

THE INFLUENCE OF SCORING METHODS UPON SCORE IN MOTOR PERSEVERATION TESTS¹

BY K. F. WALKER, R. G. STAINES AND J. C. KENNA

From the University of Sydney

I. *Theories and measures of perseveration* (pp. 51-52). II. *The construction of motor perseveration tests, and current scoring methods* (pp. 52-53). III. *The relationships between different scoring methods* (pp. 53-56). IV. *Implications of relationships between different scoring methods: (a) The distribution of individuals on a scale of 'perseveration' (pp. 56-58); (b) The theory of a 'general factor' of 'perseveration' (p. 58); (c) The theory of 'mental inertia' (pp. 58-59); (d) The diagnostic value of motor perseveration tests (p. 59). V. *Summary* (pp. 59-60). *References* (p. 60).*

I. THEORIES AND MEASURES OF PERSEVERATION

The measurement of 'perseveration' has been undertaken mainly by British psychologists, working under the stimulus of Spearman's theory of 'mental inertia'. The theory is stated in his 'law of inertia', as follows: 'Cognitive processes always both begin and cease more gradually than their (apparent) causes' ((11), p. 291). This law was enumerated on the basis of various investigations of 'perseveration', and the theoretical writings of O. Gross(4). While Neisser(9) had been the first to introduce the term 'perseveration' into psychology, it was first extensively employed by Müller & Pilzecker(8), who define it as the tendency of every idea, after having once occurred, 'to remount into consciousness spontaneously'. A few years later Gross advanced the theory of the primary and secondary function of the nervous system: 'Each nervous element whose (primary) functional excitement means the occurrence of a presentation in consciousness, persists (secondarily) after the presentation has quitted the span of consciousness. That is to say, it remains for a further long period in a state of after-function' (cited in (11), p. 44). Spearman combines the concepts of Gross and Müller: 'Joining these two concepts together, we get simply that, with some persons, there is a tendency for mental processes to persist in activity long after the cessation of the conditions to which they were originally due' ((11), p. 52).

As experimental evidence for the law of inertia Spearman quotes first the work of the Dutch school, which produced two studies in which various tests were used as measures of 'perseveration'. Only one of these tests was a motor test, in which five letters were first written forty times normally, and then twenty times in the reverse direction. A pupil of Spearman's, Wynn Jones, then administered four motor perseveration tests to seventy-seven children of about twelve years of age, and the results obtained were interpreted by Spearman as evidence of the existence of a 'general factor' of 'perseveration'. Later studies by members of the London school, notably Lankes, Hargreaves and R. B. Cattell, have produced a further body of evidence which Spearman quotes in support of his law. The tests used by these investigators to measure 'perseveration' or 'mental inertia', are almost all motor tests.

Spearman's interpretation of the results of these studies has been criticized by several

¹ Our thanks are due to Professor H. T. Lovell and Dr G. Phillips for criticism of this paper. We are also indebted to Miss N. Hamilton for making available to us the raw scores of her unpublished study, cited in the text.

52 *Influence of scoring methods upon score in motor perseveration tests*

writers(1, 6, 7) including the authors(12). This paper is not concerned with criticism of Spearman's *interpretation* of these experimental results, but with certain features of the construction and scoring of motor perseveration tests which operate to invalidate the results themselves.

II. THE CONSTRUCTION OF MOTOR PERSEVERATION TESTS AND CURRENT SCORING METHODS

Two types of motor perseveration tests have been employed, which have been called by Cattell(2), 'creative effort' tests, and 'alternation' tests, respectively. In the creative effort test the subject has to perform first a well-established activity, and then a second activity which conflicts with the first, usually one that is new to the subject. 'Perseveration' is assumed to be measured by the difference in performance of the two activities. A typical creative effort test is the 'S' test, first used by Wynn Jones. The subject has first to write as many 'S's' as possible in a given time, then to write 'Z' for the same period of time.

In the alternation test, the subject first performs one activity, then performs a second activity, and in the third part of the test is required to alternate the two. The 'S' test can be adapted into an alternation test, a typical plan of administration being as follows:

- (1) 'S' for 30 sec.
- (2) 'Z' " "
- (3) 'S' " "
- (4) 'Z' " "
- (5) 'SZ' " "
- (6) 'SZ' " "

The test is usually scored by comparing the performance of the first four parts of the test with the performance of the last two.

Both alternation tests and creative effort tests have been scored by a variety of methods. The following scoring methods have been applied to creative effort tests:

Method A, $X - Y$; Method B, $\frac{X}{Y}$; Method C, $\frac{X - Y}{X}$; where X =score on first activity, Y =score on second activity.

Method C was developed by Clarke(3) in an attempt to eliminate the influence of speed of writing upon the 'perseveration score'. For if the subject, being a rapid writer, scores high on either X or Y , or both, his 'perseveration score' will be higher than a slower writer, although he experiences no greater difficulty in changing from the first to the second activity.¹ By relating $X - Y$ to X , speed of writing is to some extent ruled out. Method B is less influenced by speed of writing than Method A, but when Rangachar(10) eliminated speed by a statistical technique, he found the distribution of his subjects' scores (calculated by Method B) to be considerably altered.

Alternation tests have been scored by the following techniques:

Method D, $\frac{X + Y}{2X}$ (where XY =score on the last 2 parts of the test, when the two activities are alternated).

Method E. (A method which eliminates interference arising from the time-allocation of the test.) Method E was first developed by Clarke(3) to provide a pure measure of difficulty in alternation. The necessity for a special method of scoring to obtain a pure measure of difficulty in alternation may be

¹ For example: Let $X=100$, $Y=90$, then $X - Y=10$. If $X=30$, $Y=20$, then $X - Y=10$. But in the first case $\frac{X - Y}{X}=10\%$; in the second case $\frac{X - Y}{X}=33\frac{1}{3}\%$.

readily demonstrated. In a test such as the 'S' test in alternation form, suppose the following scores to be recorded on the first four parts of the test:

- (1) 30 S in 30 sec.
- (2) 15 S "
- (3) 30 S "
- (4) 15 S "

The subject thus wrote 60 S in a period of 60 sec., and 30 S in another period of 60 sec. This is 1 sec. for each S, and 2 sec. for each S. When he comes to alternate S and S, each SS unit will take him $(1+2)=3$ sec. (disregarding any effects due to practice or 'perseveration'). In the fifth and sixth parts of the test he has a period totalling 60 sec., in which time he will write, at this rate, $\frac{60}{3}$ SS units, a score of 40 on XY.

His final score, calculated by Method D, $\left(\frac{X+Y}{2XY}\right)$ is $\frac{60+30}{2(40)} = \frac{90}{80}$. In this calculation we have made no allowance for any difficulty in alternation. And yet the subject is credited with an interference score without having any difficulty in alternation whatever. The greater the difference between X and Y, the smaller 2XY will be, apart from any difficulty in alternation, and the greater the 'interference score.' If the scores had been 60 S and 20 S, each SS unit would take $(1+3)=4$ sec., and the subject would do 15 SS units in a minute. 2XY would then be $2(30)=60$. Final 'interference score' would be $\frac{60+20}{60} = \frac{80}{60}$.

To obtain a pure measure of difficulty in alternation, independent of the relative difficulty (or speed) of the two activities, we used the following method (we describe our own method here rather than Clarke's, since ours is simpler):

$$\text{Interference score} = \frac{E}{A},$$

where

E = Expected score on XY.

A = Actual score on XY.

$$E = \frac{60}{\text{Time for 1S} + \text{Time for 1S}} \\ = \frac{60}{T_1 + T_2}.$$

$$T_1 = \frac{60}{\text{Score in (1)} + \text{Score in (3)}}, \quad T_2 = \frac{60}{\text{Score in (2)} + \text{Score in (4)}}.$$

$$A = \text{Score in (5)} + \text{Score in (6)}.$$

A simpler way of avoiding the influence of the initial difference in the speed of the two activities would be to specify the number of items to be done in each part of the test, and score by recording the time taken. But this method would be difficult to use in group testing. It appears, therefore, that the measure of 'perseveration' yielded by alternation tests scored by the usual method is a composite of the initial difference in ease of performance of the two activities, and the difficulty in alternating them.

III. THE RELATIONSHIPS BETWEEN THE DIFFERENT SCORING METHODS

No previous investigator has made a thorough study of the relationships between these five different scoring methods. R. B. Cattell(2) obtained a correlation of +0.52 between Methods A and D, but no other attempt to validate the indiscriminate application of creative effort and alternation tests has been made.

Two groups of data bearing upon this issue have been obtained by the present writers. The first was obtained through the courtesy of Miss N. Hamilton, who made available to us the raw scores of the battery of motor perseveration tests which she employed in an unpublished study for the Department of Psychology, the University of Sydney(5). This enabled us to rescore them by Method E, and to correlate the results with her scores, obtained by Method D. The second group of data was obtained in an investigation of our own, in which ninety-nine University students were given a battery of six motor perseveration tests, which were scored by all five methods, and the results correlated.

54 Influence of scoring methods upon score in motor perseveration tests

The tests used by Miss Hamilton were all of the alternation type, and consisted of the following:

- (1) Dot-dash.
- (2) ABC, *abc*.
- (3) S2.
- (4) 2, reversed 2.
- (5) W, reversed W.

The tests were administered to two groups of boys, ranging from thirteen to fifteen years in age, equated for I.Q., but differing in educational achievement. One, the 'normal' group ($N=50$), came from a Junior Technical School. The other, 'retarded' group ($N=75$), came from a special Activity School for educationally retarded children of normal intelligence.

The two groups showed surprising differences in the 'interference scores' yielded by Method D, the 'retarded' group being consistently higher. When the inter-test correlations¹ were computed, similar differences appeared. The intercorrelations for the two groups are given in Tables 1 and 2. The results for the 'retarded' group give clear evidence of the presence of a general factor (by tetrad analysis). The results for the 'normal' group give no evidence.

Table 1. *Inter-test correlations for 'normal' group (scoring Method D)*

Test	$N=50$				
	1	2	3	4	5
1	—	-0.07	+0.22	-0.10	+0.06
2	—	—	-0.14	+0.01	+0.05
3	—	—	—	+0.11	+0.09
4	—	—	—	—	+0.25

Table 2. *Inter-test correlations for 'retarded' group (scoring Method D)*

Test	$N=75$				
	1	2	3	4	5
1	—	+0.71	+0.72	+0.56	+0.71
2	—	—	+0.75	+0.64	+0.71
3	—	—	—	+0.55	+0.64
4	—	—	—	—	+0.78

When the tests were rescored by Method E, the inter-test correlations listed in Tables 3 and 4 were obtained. Neither matrix indicates the presence of a general factor.

Table 3. *Inter-test correlations for 'normal' group (scoring Method E)*

Test	$N=50$				
	1	2	3	4	5
1	—	0.00	+0.13	-0.20	+0.11
2	—	—	-0.16	+0.13	+0.16
3	—	—	—	-0.04	+0.02
4	—	—	—	—	+0.21

Table 4. *Inter-test correlations for 'retarded' group (scoring Method E)*

Test	$N=75$				
	1	2	3	4	5
1	—	+0.31	+0.37	-0.04	+0.04
2	—	—	+0.27	-0.19	+0.08
3	—	—	—	+0.14	+0.14
4	—	—	—	—	+0.23

¹ 'Product moment' correlations ($r = \sqrt{\frac{\sum xy}{n\sigma_x\sigma_y}}$) were calculated throughout.

The correlations between Methods D and E were calculated, and are listed in Table 5. The striking difference between the two groups is maintained, the correlations being high for the 'normal' group and negligible for the 'retarded' group.

Table 5. *Correlations between scoring Methods D and E*

	'Normal' group	'Retarded' group
Test 1	+0.69	+0.01
2	+0.67	-0.19
3	+0.92	+0.06
4	+0.68	+0.11
5	+0.73	+0.01

In our own investigation the following battery of tests was used, all the tests being of the alternation type:

- (1) ABC *abc*.
- (2) $\pm$ ∞ .
- (3) $\boxplus$ H.
- (4) $\mathfrak{d}$ 5.
- (5) $\sqcap$ $\sqcap$.
- (6) £ Z.

Tests 4, 5, and 6 had not been previously employed by any investigator. It will be noted that the tests fall into three groups:

- (a) Those containing *two old* activities (Tests 1 and 6).
- (b) Those containing *two new* activities (Tests 2 and 5).
- (c) Those containing one old and one new activity (Tests 3 and 4).

In group (c) the new activity was given *first*.

Each test was scored by the five methods described above, and the correlations between each method and every other method were calculated. These correlations are listed in Table 6.

Table 6. *Correlations between various scoring methods on six motor perseveration tests administered to a group of ninety-nine male university students*

Correlation between Methods	Test 1 ABC, <i>abc</i>	Test 2 $\pm$ ∞	Test 3 $\boxplus$ H	Test 4 $\mathfrak{d}$ 5	Test 5 $\sqcap$ $\sqcap$	Test 6 £ Z
A & C	+0.88 ¹	+0.93 ¹	+0.57 ¹	+0.45 ¹	+0.25 ¹	+0.04
A & B	+0.89 ¹	+0.97 ¹	+0.76 ¹	+0.67 ¹	+0.96 ¹	+0.62 ¹
A & D	+0.16	+0.34	-0.01	+0.20	+0.10	-0.48 ¹
A & E	+0.14	+0.41 ¹	+0.05	+0.39 ¹	-0.35 ¹	-0.43 ¹
A without sign & B	+0.53 ¹	+0.19	-0.80 ¹	-0.64 ¹	-0.08	-0.77 ¹
A without sign & C	+0.41 ¹	+0.21	-0.34 ¹	-0.51 ¹	-0.19	-0.30 ¹
A without sign & D	+0.22	+0.41 ¹	-0.07	-0.02	+0.09	+0.45 ¹
A without sign & E	+0.06	+0.03	-0.16	-0.21	+0.02	+0.28 ¹
B & C	+0.85 ¹	+0.81 ¹	+0.33 ¹	+0.50 ¹	+0.71 ¹	+0.94 ¹
B & D	+0.03	+0.35 ¹	+0.07	+0.26	+0.01	-0.96 ¹
B & E	+0.16	+0.34 ¹	+0.01	+0.44 ¹	-0.05	-0.36 ¹
C & D	+0.15	+0.31 ¹	+0.37 ¹	+0.20	+0.08	-0.29 ¹
C & E	+0.08	+0.34 ¹	+0.15	+0.08	+0.03	-0.14
D & E	+0.85 ¹	+0.85 ¹	+0.68 ¹	+0.69 ¹	+0.90 ¹	+0.65 ¹

Note. Probable errors range from 0.004 to 0.069.

¹ Indicates a statistically significant correlation ($r > 4$ P.E.).

It will be noted that Methods A, B, and C inter-correlate highly on each test; that the correlation between Methods D and E is higher on some tests than on others; that the correlations of Method D with Methods A, B and C are all low or negative and the same is true of correlations between these methods and Method E.

IV. IMPLICATIONS OF RELATIONSHIPS BETWEEN DIFFERENT SCORING METHODS

(a) *The distribution of individuals on a scale of 'perseveration'*

Cattell obtained a relatively high positive correlation between Methods A and D, but our correlations are all either low or negative. The reason for this is that if $X-Y$ is large, irrespective of its algebraic sign, XY will be small, whether or no any difficulty in alternation is experienced. This follows from the construction of the alternation test, as we demonstrated in the example previously given.¹ A large initial difference between the two activities therefore gives a high 'interference score' by Method D, $\frac{X+Y}{2XY}$, whether the initial difference be positive or negative. But if $X-Y$ is negative, instead of the first activity 'perseverating' and interfering with the second, it apparently has a facilitative effect. An individual who returns a negative $X-Y$ score is usually classified as a Low Perseverator. In our tests many negative $X-Y$ scores were returned, and this accounts for our correlations between Methods A and D being either insignificant or negative. An individual with a large negative $X-Y$ score appears as a Low Perseverator by Method A, yet returns a high 'interference score' by Method D. The distribution of a group of individuals will thus be very different according to the scoring method used.

This analysis holds for the other methods of scoring creative effort tests (Methods B and C), and whenever negative scores are recorded by these methods, results will differ materially according to which type of test is employed.

Cattell's tests all yielded positive $X-Y$ values, and this is what gave him positive correlations. If our tests are re-scored by reversing the $X-Y$ values, the negative correlations become positive, and Test 6, which gave large (but negative) $X-Y$ values by Method A yields the significantly high correlations of +0.48, +0.29 and +0.96, for r_{AD} , r_{BD} and r_{CD} respectively.

If the correlation between the creative effort methods of scoring and Method D is determined by the size of $X-Y$, irrespective of its sign, it follows that the correlation will be small in tests which yield low $X-Y$ values, irrespective of sign. Cattell's tests all yielded high $X-Y$ scores, since they involved one old and one new activity. The only test in our battery which gave large $X-Y$ values is Test 6. This test contained two old activities, but one ('Z') can be written much faster than the other ('L'). And since the slower one was given first, the $X-Y$ values, though high, are negative. As a result the correlation between Method D and the creative effort methods for this test is high and negative.

This explains the discrepancy between our results and those of Cattell, and suggests that whenever a test yields low or negative $X-Y$ values, it is of considerable importance whether we score it as a creative effort test or an alternation test.

¹ The example of p. 53 had a positive value for $X-Y$, but a negative $X-Y$ value gives the same result. Let $X=30$, $Y=60$. Each XY unit then takes $(2+1)$ sec.; and in 60 sec. the subject will do $\frac{60}{3} XY$ units, giving an interference score by Method D of $\frac{30+60}{2(40)} = \frac{90}{80}$, as before.

Turning now to the different methods of scoring alternation tests, the correlations of Methods D and E may be compared. In Miss Hamilton's results they vary according to the group considered; in our results they vary according to the test considered.

In Miss Hamilton's results the correlations of Methods D and E are high and positive for the 'normal' group, but are negligible for the 'retarded' group. The 'normal' group yielded consistently low 'interference scores' by Method D; the 'retarded' group recorded high 'interference scores' by this method. Method E eliminates the influence of initial difference upon 'interference score'. Since Methods D and E show insignificant correlations for the 'retarded' group, it appears that this group's scores by Method D are indicative of large initial differences in their speed on the two activities rather than of any difficulty in alternation. Method D does not distribute individuals on a scale of 'alternation interference', then, but predominantly in terms of the initial difference in speed of the two activities. Whenever this initial difference is large, considerable error is introduced by employing Method D, if it is desired to measure 'alternation interference'.

This conclusion is supported by the correlations of Methods D and E in our own results. For Tests 1, 2, and 5 the correlations are very high and positive. These are all tests which show little initial difference in speed of the two activities, containing either two old or two new activities. But Test 6 also contains two old activities, yet it gives a much lower correlation. This is because it takes longer to write '£' than 'Z', in spite of equal familiarity with both symbols. Tests 3 and 4 each contain one old and one new activity, and yield high $X-Y$ values as a result (negative ones since the new activity preceded the old). These tests likewise give lower correlations between the two methods. It would appear, then, that the greater the initial difference in the speed of the two activities, the greater the difference between the distribution of a group in terms of Method D and a distribution of the same group by Method E.

The assumption of the London school is that 'alternation interference' is a measure of 'perseveration', as defined by Spearman. They have attempted to measure 'alternation interference' only in tests showing large initial differences between the two activities, employing only Method D. This means that their classification of individuals as High and Low Perseverators is a classification largely in terms of High and Low Initial Difference. The correlation between 'alternation interference' (given by Method E) and initial difference (given by Method A without regard to sign) is shown in Table 6, and is negligible for each test. Pinard's classification of his group of children into High, Low and Moderate Perseverators was thus a classification in terms of their initial difference in performance of the two activities in his tests, and not a classification based on their relative difficulty in alternation at all. The association which he found to exist between certain traits of personality and scores on alternation tests may not have occurred in a group which recorded low or negative $X-Y$ scores (since the distribution of individuals would then have been significantly different), or if he had used tests which yielded low or negative $X-Y$ values because they contained two old or two new activities rather than one old and one new.

The position is, then, that if the initial difference is large and positive, it matters little whether we use the test in its creative effort form or as an alternation test (i.e. Methods A and D have high positive inter-correlation). But the larger the initial difference, the lower the correlation between Methods D and E, so that if the initial difference is large, and we use the test in its alternation form, scoring by the usual Method D, the scores

we obtain will not indicate mere difficulty in alternation, but mainly the size of the initial difference. It would appear, then, that there is no point in using alternation tests unless they are scored by Method E. When they are scored by this method, it appears that ease or difficulty in alternation bears no relation to the size of the initial difference, that is, to the ease or difficulty of 'creative effort'. The association which the London school claims to exist between them is a spurious association resulting from their method of constructing and scoring motor perseveration tests.

(b) *The theory of a 'general factor' of 'perseveration'*

This conclusion has important bearing on Spearman's claim that the work of the London school has proved the existence of a general factor participating in all 'perseverative' processes. We have criticized the logical and statistical evidence for this claim elsewhere (12). The present conclusion, however, suggests, even if there were no flaws in the statistical evidence, that the general factor which Spearman claims to have found cannot be one of 'perseveration' as he defines it. For it appears, contrary to his claims, that ease or difficulty in alternation (when accurately measured) is not related to ease or difficulty in 'creative effort' at all. (As noted above, the correlations of initial difference (Method A without sign) and 'alternation interference' (Method E) are all low.)

This conclusion is supported by the changes which occur in Miss Hamilton's inter-test correlations when they are re-scored by Method E. By Method D the inter-test correlations for the 'retarded' group were high, and gave evidence of the presence of a general factor. When re-scored by Method E, however, this evidence disappears. When the influence of large initial difference is removed, then, the evidence for a general factor disappears. It will be remembered that the 'normal' group did not give evidence of a general factor. The inter-test correlations for this group are not significantly changed when the influence of initial difference is removed. This is not surprising, for this group showed little initial difference.

Altogether, then, it appears that a group factor which appears in a battery of alternation tests scored in the usual way is a factor of 'creative effort', if anything. It results from the fact that all the tests contain one old and one new activity.

(c) *The theory of 'mental inertia'*

The cornerstone of Spearman's theory of 'mental inertia' is the proof that a general factor of 'lag' exists, functioning in all cognitive processes. Whether such a factor exists is very doubtful indeed, and if it does, it cannot be a factor of 'mental inertia' as defined by Spearman. The concept of 'mental inertia' refers only to the lag of one process so that it interferes with the performance of another activity initiated during the period of the lag. But the only general factor for which there is any evidence at all is a factor of 'creative effort', i.e. the difficulty in forming a new habit contrary to another well-established one, presumably due to the rigidity of the disposition founded by the exercise of the old-established habit. There is no reason to suppose that disposition-rigidity exerts a disruptive influence upon the performance of a new, conflicting activity only when the new activity is initiated immediately after the performance of the old activity. One has more difficulty in writing Z than in writing S *at any time*, whether or not one has just been writing S. This is shown clearly in our results, which record high negative X-Y values, in spite of the fact that the new activity preceded the old. If the difficulty of 'creative

effort' results solely from the temporal succession of the old and the new activities, why should our $X-Y$ value be negative?

It would appear that Spearman has attempted to verify his hypothesis of 'mental inertia' by experimental evidence concerning phenomena which he subsumes under his "law of retentivity of dispositions", which states "cognitive events by occurring establish dispositions which facilitate their recurrence" (11, p. 271). He is careful to distinguish these phenomena from those of 'perseveration' in theory, but in practice they have been mingled indiscriminately. He writes, in explanation of "the seeming paradox that the persistence involved in perseveration is so widely unlike that involved in dispositions", that "naturally enough, the power of switching the energy from one to another system of neurones has nothing to do with the susceptibility of these neurones to retain engrams afterwards" (11, p. 307). But his measurement of 'perseveration' is confined to instances in which the power of switching from one to another system of neurones is largely dependent upon the relative rigidity of the two systems. When the influence of this rigidity is experimentally eliminated, the evidence for a general factor of 'perseveration' disappears, and with it goes the evidence for an inertia manifested in all cognitive processes.

(d) *The diagnostic value of motor perseveration tests*

Motor perseveration tests have been employed, notably by Pinard (9a), Stephenson (11a, 11b, 11c) and Cattell (2), to diagnose the presence of certain personality-traits in children and adults. Our results do not impugn the validity of such diagnosis. They indicate rather that the explanation of these traits in terms of individual differences in 'mental inertia' is inadmissible. The explanation appears to be that these traits are associated with individual differences in difficulty of making 'creative effort', and the cause of this association is something that remains to be finally demonstrated. A step in this direction has been made in the work of Stephenson (11a, 11b, 11c). Discussion of this problem is beyond the scope of this paper, and will be included in another study at present in preparation.

V. SUMMARY

1. Spearman's theory of 'mental inertia' has been criticized on various grounds. In this paper certain invalidities arising from the methods of construction and scoring motor perseveration tests are discussed.

2. Motor perseveration tests are of two types: tests of 'creative effort', and 'alternation tests'. Five methods of scoring have been employed, but the relationship between them has not been investigated.

3. Data are reported from two investigations, in which batteries of motor perseveration tests were administered to several groups, and the relationships between five possible methods of scoring determined by correlation.

4. The implications of these data are:

(a) that the measure of 'perseveration' yielded by alternation tests scored by the usual method is a composite of the initial difference in ease of performance of the two activities, and the difficulty experienced in alternating them;

(b) that alternation tests will give the same distribution of individuals as creative effort tests only when an old and a new activity are involved in the test;

(c) that the initial difference in ease of performance of the two activities bears no relation to the difficulty experienced in alternating them;

(d) that, as a result, any 'general factor' which may be yielded by a test battery which includes alternation tests is a factor of initial difference in ease of performance of the two activities;

(e) that initial difference in ease of performance of two activities is not entirely a matter of temporal sequence, and therefore not a pure measure of 'mental inertia';

(f) that the personality traits which have been found to be associated with scores on 'perseveration' tests are not capable of explanation in terms of individual differences in 'mental inertia'.

REFERENCES

- (1) BURRI, C. (1935). "The present status of the problem of individual differences in alternating activities." *Psychol. Bull.* xxxii, 113-39.
- (2) CATTELL, R. B. (1935). "On the measurement of perseveration." *Brit. J. Educ. Psychol.* v, 76.
- (3) CLARKE, G. (1934). "Some character traits of delinquent and normal children in terms of perseveration factor." *Aust. Council Educ. Research Series*, No. 29.
- (4) GROSS, O. (1902). *Die cerebrale Sekundärfunktion*. Leipzig. Pp. 69.
- (5) HAMILTON, N. "Perseveration and stability in school children." *Unpublished Study*, University of Sydney Psychology Department Library.
- (6) JASPER, H. H. (1931). "Is perseveration a functional unit participating in all behaviour processes?" *J. Soc. Psychol.* ii, 28-51.
- (7) KELLEY, T. (1928). *Crossroads in the Mind of Man*. Stanford University Press.
- (8) MÜLLER, G. E. and PILZECKER, A. (1900). "Experimentelle Beiträge zur Lehre vom Gedächtniss." *Z. Psychol. Ergbd.* i. Pp. xiv+300.
- (9) NEISSER, C. (1894). "65 Sitzung des Vereins Ostdeutscher Irrenärzte." *Allg. Z. Psychiat.* li, 1016.
- (9a) PINARD, J. W. (1932). "Tests of perseveration. I. Their relation to character." *Brit. J. Psychol.* xxiii, 5-19.
- (10) RANGACHAR, C. (1932). "Differences in perseveration among Jewish and English boys." *Brit. J. Educ. Psychol.* ii, 199-211.
- (11) SPEARMAN, C. (1927). *The Abilities of Man*. London, Macmillan.
- (11a) STEPHENSON, W. (1931). "Studies in experimental psychiatry. I." *J. Ment. Sci.* lxxvii, 723-41.
- (11b) STEPHENSON, W. (1932). "Studies in experimental psychiatry. II." *J. Ment. Sci.* lxxviii, 315-50.
- (11c) STEPHENSON, W. (1932). "Studies in experimental psychiatry. III." *J. Ment. Sci.* lxxviii, 908-28.
- (12) WALKER, K. F., STAINES, R. G. and KENNA, J. C. (1941). "Is there a general factor of perseveration?" *Aust. J. Psychol. Phil.* xix, 58-75.

(Manuscript received 25 November 1943)

ON IDENTIFYING AESTHETIC TYPES

BY E. A. PEEL

(1) *The problem* (pp. 61–62). (2) *Tests and population* (pp. 62–64). (3) *Results and conclusions:* (a) *Landscape test* (pp. 64–66), (b) *Still-life test* (p. 67). (4) *Summary* (p. 67). (5) *Mathematical Appendix* (pp. 68–69). *References* (p. 69).

(1) THE PROBLEM

Previous research by Burt⁽¹⁾, Dewar⁽²⁾ and Eysenck^(3, 4) has shown the existence of a general and bipolar factor in aesthetic judgements. It was realized that the general factor might be of a composite nature, and it was also identified with a sense of artistic form and good taste. The possibility that irrelevant factors might enter into the judgements of non-experts has been considered by Burt⁽¹⁾. The bipolar factors have been likened to Bullough's types⁽⁵⁾. Eysenck⁽⁴⁾ is led to suspect that the bipolar factor relates to preference for brightness-restraint of colour and content and thence, after constructing a test to measure this property, finds that it correlates positively with temperamental tests of extraversion-introversion, radicalism-conservatism and youth-age. Thus, where identification of aesthetic factors has been attempted, it has been with reference to temperamental traits, and the emphasis has been on the person rather than the artistic qualities of the picture.

It is the object of this paper to outline a method for comparing aesthetic preferences with these artistic qualities and for obtaining an estimate of a person's aesthetic preference in terms of them.

We begin by considering a test composed of a number of pictures which are to be arranged according to the order in which they are 'liked' or preferred aesthetically. This arrangement of pictures beginning with the picture most liked will be known as the order of 'liking'. In detail each person's order of 'liking' would be different, but we might expect certain broad similarities depending upon the nature of the test and the group of persons tested. The orders may be compared by correlation and the correlation coefficients analysed to give the factors which characterize the group of individuals. The aesthetic qualities related to the factors may be conjectured by an examination of the pictures which comprise the test. Supposing, for example, that a certain group of individuals arranged a collection of pictures in such a way that those by Monet, Pissarro, Sisley and Constable were placed at the top and those by the English Victorians and Pre-Raphaelites at the bottom, the possibility that the general factor indicated a preference for 'Impressionism' would have to be considered when interpreting the results. The problem is whether we can make this guess more explicit and exact by making a comparison which would give us a quantitative estimate of a preference for impressionist painting.

We can obtain such a measure by asking a team of experts to arrange the items *not* in order of 'liking' but in order of 'Impressionism', that is to say, by choosing first that item which is the most impressionistic and proceeding so on through the test. If 'Impressionism' is a more or less objective quality of the pictures, continuously distributed throughout the test, we should expect fairly close agreement between different experts and we could average their orders to obtain an order of 'Impressionism' for the pictures. We have now the means of making a quantitative comparison between a person's order

of 'liking' and the order of the pictures according to 'Impressionism'. This would be done by correlating the two orders item by item, and those persons whose 'liking' correlated highly with 'Impressionism' would be said to have a preference for 'Impressionist' painting. One such objective order as 'Impressionism' might not however exhaust either the entire artistic qualities of the pictures making up the test, or the complete grounds of aesthetic preference in the group of persons tested. The external team of experts might then be required to arrange the pictures in order of, say, 'realism', 'quality of technique', 'composition', 'colour' and so on. Each person's order of 'liking' could then be correlated with each of the orders according to these qualities or *criteria* as they will be called, and an estimate of any person's aesthetic choice obtained in their terms, thus giving a more complete description of his 'liking'. An advantage in using this method of analysis is that the person is quite unaware of the process or even intention to analyse his aesthetic preferences.

For any given test there would be as many orders of 'liking' as persons tested and as many orders according to 'criteria' as artistic qualities selected from the pictures. By correlation of these orders we may obtain three different matrices of correlation coefficients:

- (a) Correlations between the orders of 'liking' for the persons.
- (b) Correlations between the orders of 'liking' and the criteria.
- (c) Correlations between the criteria.

Using matrix notation these matrices may be combined to give a single matrix R , where

$$R = \begin{bmatrix} R_L & R_O \\ R'_O & R_C \end{bmatrix},$$

and R_L denotes the matrix of correlations between the orders of 'liking' for the persons, R_O the matrix of correlations between the orders of 'liking' and the criteria, and R_C the matrix of correlations between the criteria.

Three interesting statistical possibilities are suggested by the matrix of correlation coefficients R . In the first place it would be possible to analyse the entire matrix R and so obtain general and bipolar factors, entering into the correlations between both persons and criteria combined. The axes could then be rotated about the vectors corresponding to the criteria and from inspection of the rotated loadings the influence of the criteria in the group of persons' aesthetic choice could be assessed.¹ It is evident at this stage that criteria are preferable which do not correlate significantly amongst themselves. Secondly, the regression estimate of each person's 'liking' on the artistic qualities or criteria could be obtained, and in this way the influence of each criterion on each person could be estimated and the extent to which the criteria account for each person's 'liking' could be assessed. Lastly, by first analysing the submatrix of 'liking' inter-correlations (R_L), and then obtaining the estimate of these factors on the criteria, it is possible to identify the factors characterizing the whole group of persons in terms of the criteria selected.

(2) TESTS AND POPULATION

The method of analysis has been applied to several tests of abstract and representational painting and drawing and from these tests, two are selected for description. The first, a

¹ The results of applying this method of factorial analysis to tests of art appreciation were discussed by the writer in a paper read to the Scottish Section of the British Psychological Society at Edinburgh in November 1943. Some mathematical and statistical aspects of the method, and a more complete consideration of the statistical properties of ideal criteria, are discussed in an unpublished thesis.

landscape test, consists of thirty-one post-card size coloured reproductions of landscapes painted by a wide range of artists from the Middle Ages to modern times. The theme is confined to landscape in order to restrict as far as possible any irrelevant factors which might be introduced by diverse subject matter and content.

The three qualities finally selected as criteria for use in conjunction with this test were 'detailed naturalism', 'composition' and 'atmospheric light'.

'Detailed naturalism' is used to denote the detailed imitation or description of nature. It results from looking at nature by means of what has been called 'Consecutive vision' ((6), pp. 245-8). The artist turns his attention to each detail of the scene as he comes to paint it—a tendency noted among the Pre-Raphaelites and characteristic of the Victorian 'genre' art. Good examples of the technique are seen in Millais' *Ophelia* and Frith's *Derby Day*.

By 'composition' is meant the unified arrangement in space of the elements of a picture. These elements are composed of lines, masses, vista, suggested movement, trees, clouds, figures, lights, shadows, and colours, etc. A good composition is one in which the elements cohere, balance, and preserve a satisfying harmony and unity.

The artistic criterion 'light' is taken to mean the general atmospheric effect of light produced in a painting as opposed to sombre landscape.

The second test is composed of thirty-one still-life paintings. A still-life group is set up, consisting of a coffee pot, a glass decanter, a cloth, a plate containing cheese and tomatoes, and two books. Water colour drawings of size $10\frac{1}{2} \times 7\frac{1}{2}$ in. are made of the group, some of which are coloured and some uncoloured. This test includes drawings which are mere reproductions, others which are intellectual interpretations and abstractions, and others in which the group forms the merest symbol or release for an artistic expression. The drawings show varying degrees of draughtsmanship apart from any resemblance to the original group. In a few items only is the composition or arrangement of the group considerably changed.

In this test there appears to be a varied collection of pictures showing possibilities of a wide choice of criteria and at the same time the content has been maintained constant, thus eliminating any variations due to association.

The three criteria finally retained for this test were:

(i) Realism, which implies the extent to which the pictures are good reproductions of the original group. Since it was impossible to exhibit this group to the judges for comparison, one picture was selected as the best copy and the rest of the items ranked in order of similarity to this picture.

(ii) Technical quality, which refers to draughtsmanship and colour washing only, apart from any resemblance of the drawing to the original group.

(iii) Spontaneity, which indicates the freedom of expression, the absence of laboured efforts, and evidence of the 'dash' which may be considered a characteristic of a successful sketch or impression.¹

¹ An artist criticized the drawings on the grounds of sincerity; he contended that drawings which represent such a wide scale of artistic creation could hardly be equally sincere. The question of sincerity may be closely related to that of spontaneity, but whereas an order seems possible for spontaneity by judging the quality of the lines and washes, an order for sincerity would be very difficult to assess in relation to paintings designed for test purposes.

The persons who were selected as subjects for the investigation included art specialists, and non-expert adults. They were all required to order the items of each test in order of 'liking', beginning with the item most liked. The team of expert judges was drawn from artists, art teachers and postgraduate art students. They were required to arrange the items according to the criteria or artistic qualities. If their average orders for the criteria complied with the requirements of the statistical tests devised (see the Appendix), then the subsequent use of the criteria for comparison with the orders of 'liking' was considered to be justified.

(3) RESULTS AND CONCLUSIONS

The complete analysis of the matrix of correlations R , involving the calculation of the regression estimates of the persons' 'liking' on the criteria and the estimates of the factors which characterize the group of persons, has been carried out only for the landscape test.

(a) *Landscape test.* Twelve persons arranged the items in order of 'liking' and the factors obtained by analysis of their correlations, using the centroid method, are given in Table 1.

Table 1

Person	<i>Decimal points are omitted</i>			
	Factor I		Factor II	
	+	-	+	-
A	.	665	.	491
B	.	552	.	271
C	.	389	.	319
D	.	286	.	574
E	608	.	.	406
F	740	.	.	286
G	532	.	101	.
H	414	.	.	088
J	.	048	.	093
K	150	.	.	692
L	.	477	.	132
M	573	.	.	464

Significance of the factors: 34% of correlation coefficients exceed twice their standard error, 6.7 % of first residuals exceed twice the standard error of parent correlation coefficients.

There is thus evidence of significant first and second factors.

We next consider how each person's 'liking' correlates with the three criteria. The correlation coefficients are given in Table 2.

Table 2

Person	Criteria		
	Detailed naturalism	Composition	Atmospheric light
A	-473	473	-198
B	-385	324	-203
C	-324	549	-006
D	-066	670	-099
E	637	055	-352
F	720	033	-280
G	456	011	-088
H	297	187	-302
J	-220	225	093
K	143	235	-220
L	-396	-017	-154
M	429	253	-115

The regression coefficients are then obtained and by means of these coefficients any person's 'liking' can be estimated as a weighted sum of the criteria.

Table 3. *Table of regression coefficients*

Person	Criteria			Max. corr.
	Detailed naturalism	Composition	Atmospheric light	
A	-497	422	-332	707
B	-423	283	-314	564
C	-280	517	-092	613
D	-011	674	-127	761
E	604	140	-214	684
F	708	127	-111	740
G	470	070	020	462
H	267	230	-248	444
J	-185	200	042	300
K	141	361	-200	430
L	-466	-065	-261	475
M	464	312	-018	529

Of these regression coefficients 16.7 % exceed $3.0 \times$ standard error; 33.3 % exceed $2.0 \times$ standard error; 52.8 % exceed $1.5 \times$ standard error; using the formula described by Thomson ((9), p. 16).

The data of Table 3 could be expressed alternately in the form of regression equations. Thus for the 'liking' of persons A, D and E we have:

Estimated liking of	Detailed naturalism	Composition	Atmospheric light
A	-0.497	+0.422	-0.332
D	-0.011	+0.674	-0.127
E	0.604	+0.140	-0.214

It is possible to obtain not only the regression of each person's 'liking' on the criteria but also the regression estimates of the factors which characterize the 'liking' of the entire group of persons tested. This may be done by a method outlined in the Appendix, and when it is applied to the above data the estimates shown in Table 4 are obtained for the two factors of Table 1.

Table 4. *Factor estimates*

Factor	Detailed naturalism	Composition	Atmospheric light
I	566	-070	062
II	088	505	-272

Upon rotation about 'detailed naturalism' the following coefficients are obtained.

Table 5

Rotated factor	Detailed naturalism	Composition	Atmospheric light
I	573	010	019
II	000	489	-278

In considering the data presented in Tables 1-5 we may first note the bipolar nature of the first factor extracted in Table 1. It is followed by the second factor which is virtually general. The group of individuals A, B, C and D, comprised of persons with artistic training and experience, possesses loadings in the bipolar factor which have a common sign. From Tables 2 and 3 it is evident that the bipolar character may be closely related to a preference for or dislike of landscape paintings executed in a detailed naturalistic manner. The group A, B, C and D would seem, by their negative correlations with 'detailed naturalism' in Table 2 and their negative coefficients for 'detailed naturalism' in

Table 3, not to show preference for this quality. The non-expert adults on the other hand, with the exception of K and L, appear to be characterized by a preference for detailed reproduction of nature, as inspection of Tables 2 and 3 reveals.

The general factor of Table 1 might be interpreted as revealing the response of the persons to 'composition'. This possibility is confirmed by the positive correlations with 'composition' in Table 2 and its positive coefficients in Table 3. It is to be noted that the regression coefficients for the artistic group A, B, C and D with respect to 'composition' are generally more significant than the coefficients for the non-experts. That the bipolar factor probably corresponds to a response to 'detailed naturalism' and the general factor to a response to 'composition' is clearly supported by the regression estimates of these factors, given in Tables 4 and 5. Upon rotating the factors about 'detailed naturalism' it is seen that the first factor is very closely associated with a response to this artistic criterion, whilst the principal influence in the case of the second factor appears to be 'composition'. The marked influence of 'detailed naturalism' and 'composition' on these factors is also supported by the alternative method of analysis into its factors of the entire matrix *R*, inclusive of both persons and criteria, which was referred to in the first footnote. Lack of space forbids that more information can be given than the factor loadings which are obtained if this is done. These loadings are shown in Table 6.

Table 6

Adult	Centroid loading				Rotated loading			
	I		II		I		II	
	+	-	+	-	+	-	+	-
A	.	650	397	.	.	547	526	.
B	.	515	189	.	.	461	298	.
C	.	434	453	.	.	324	538	.
D	.	338	695	.	.	177	752	.
E	633	.	198	.	660	.	054	.
F	740	.	258	.	780	.	089	.
G	508	.	128	.	524	.	013	.
H	380	.	243	.	424	.	154	.
J	.	123	175	.	.	082	198	.
K	124	.	627	.	257	.	584	.
L	.	426	.	109	.	439	.	013
M	494	.	501	.	593	.	379	.
Detailed naturalism	882	.	199	.	904	.	000	.
Composition	.	311	805	.	.	127	854	.
Atmospheric light	.	210	.	329	.	278	.	275

Factor I is rotated about 'detailed naturalism'. The identification of these factors with 'detailed naturalism' and 'composition' is again clearly confirmed. The orthogonal character of these artistic criteria is also to be noted. Such criteria are preferable for analysis purposes.

The influence of 'atmospheric light' does not appear to be so marked, and this opinion is confirmed generally by inspection of its correlation and regression coefficients in Tables 2 and 3.

Within the limitations imposed by the statistical and psychological significance of the data, we may conclude that non-expert adults have a marked preference for naturalistic landscape paintings and are influenced to a less extent by 'good composition'. Artists, on the other hand, reveal a definite preference for 'good composition' and tend if anything to prefer landscape paintings which are not so detailed in their representation.

(b) *Still-life test.* This test was applied to the group of persons who arranged the items of the landscape test and the results obtained confirmed its findings. The still-life test, however, proved to be an inferior test in that its items were too heterogeneous, ranging from close reproduction to remote abstraction. Such large differences between the items tend to promote considerable similarity of choice between the persons' orders of 'liking'. It is also difficult to find a criterion which is uncorrelated with the three criteria, 'realism', 'technique', and 'spontaneity'. These criteria correlate significantly amongst themselves:

Realism-Technique	0.555
Technique-Spontaneity	-0.802
Spontaneity-Realism	-0.508

'Realism' and 'Technique' therefore overlap to a large extent and 'Spontaneity' is bipolar to them both, as is shown in the first column of factor loadings in Table 7.

Table 7. *The results obtained by analysing the complete matrix R, inclusive of persons and criteria*

	Centroid loading			
	I		II	
	+	-	+	-
Adult				
A	.	047	.	670
B	249	.	.	468
C	409	.	.	555
D	695	.	.	082
E	966	.	162	.
F	961	.	006	.
G	918	.	082	.
H	792	.	.	036
J	815	.	238	.
K	840	.	056	.
L	814	.	.	217
M	878	.	.	231
Realism	749	.	.	252
Technique	892	.	226	.
Spontaneity	.	704	.	458

Significance: correlations, 73 %, exceed 2 standard error; residuals, 5.7 %, exceed 2 standard error of parent correlation.

From the first factor it may be observed that the persons A, B and C are less influenced by 'realism' and 'technique' than the remaining adults, whilst from the second factor it appears that A, B and C respond more favourably to 'spontaneity' of drawing.

(4) SUMMARY

The selection of an optimum set of artistic criteria, and the subsequent comparison of the persons' orders of aesthetic preference with these criteria by means of correlation, provides a basis of analysis for calculating the estimate of the persons' 'liking' in terms of the artistic qualities of the items, and for analysing the correlations into factors characterizing the group of persons and criteria. By either of these approaches it is possible to assess the factors which characterize the attitude of a group of persons towards any test of appreciation, in terms of the artistic qualities of the test items. The method emphasizes the qualities of the work of art rather than the temperamental qualities of the person, but its results are complementary to those of earlier researches in which the emphasis was placed on aesthetic types.

MATHEMATICAL APPENDIX

(a) *Item ranking.* The items were ranked for both orders of 'liking' and orders according to the criteria in eleven classes in which the frequencies approximated to normal type:

Score	-5	-4	-3	-2	-1	0	1	2	3	4	5
Frequency	1	2	2	3	4	7	4	3	2	2	1
Least liked.....											Most liked
Least of a quality											Most of a quality

It was thus possible to use the product moment correlation coefficient.

(b) *Validity of the criteria.* The technique evolved for testing the validity of the criterion rankings of the test items takes place in four stages.

(i) The item rankings are obtained by different judges according to the selected criterion.

(ii) The variance of the rankings obtained by the group of judges is analysed to determine how significantly the variance within each item compares with that between item means.

(iii) If the variance within items is not significant, the average ranking for each item is obtained.

(iv) If there were perfect agreement between the judges, the distribution of the average rankings would be the same as that quoted in Section (a) of this appendix. The smaller the agreement between the judges, the greater is the tendency for the distribution of average ranks to cluster about a central class near to zero. The degree to which the actual distribution of average ranks approximates to the perfect case is measured by using the χ^2 test of goodness of fit, and if this test is satisfied, the average items ranks are rescaled to the ideal distribution.

The ultimate implication in the process of rescaling is that the use of the rescaled distribution does not seriously change the value of the resulting correlation coefficients obtained by its substitution.

The probabilities derived for the χ^2 test exceeded 20 % for each of the criteria used. The distribution of the average order was therefore considered sufficiently close to the perfect distribution to justify a rescaling.

(c) *Estimate of the person factors on the criteria.* We consider the matrix

$$R = \begin{bmatrix} R_L & R_O \\ R'_O & R_C \end{bmatrix},$$

and analyse the submatrix R_L into its factors to give the person specification equations

$$z = M_L f. \quad (1)$$

The regression estimate of one person's 'liking' on the criteria is given by

$$\hat{z}_S = r'_O R_C^{-1} c,$$

where r'_O is the row vector of the correlations between his 'liking' and the criteria. Then for the column vector of persons

$$\hat{z} = R_O R_C^{-1} c. \quad (2)$$

After Thomson ((7), p. 306 and (8), p. 41) we have for the estimate of the factors

$$\hat{f} = M'_L R_L^{-1} z, \quad (3)$$

whence by substituting $\hat{z}$ for z by equation (2), we obtain the estimate

$$\hat{\hat{f}} = M'_L R_L^{-1} R_O R_C^{-1} c. \quad (4)$$

This estimate of f is really an estimate based on an estimate.

Equation (4) gives an estimate of all the factors including the specifics. In this research we require only the estimate of the common factors f_O which is given by the equation

$$\hat{\hat{f}}_O = M'_{LO} R_L^{-1} R_O R_C^{-1} c.$$

The matrix multiplier $M'_{LO} R_L^{-1}$ may be computed by using the method due to Ledermann (see Thomson (7), p. 110).

The writer wishes to acknowledge the help and interest of Professor Godfrey H. Thomson in this paper.

REFERENCES

- (1) BURT, C. (1940). *The Factors of Mind*, chapter VI. Univ. London Press.
- (2) DEWAR, H. (1938). "A comparison of tests of artistic-appreciation." *Brit. J. Educ. Psychol.* VIII, 29-49.
- (3) EYSENCK, H. G. (1940). "The general factor in aesthetic judgements." *Brit. J. Psychol.* XXXI, 94-102.
- (4) EYSENCK, H. G. (1941). "Type factors in aesthetic judgements." *Brit. J. Psychol.* XXXI, 263-70.
- (5) BULLOUGH, E. (1908). "The perceptive problem in the aesthetic appreciation of single colours." *Brit. J. Psychol.* II, 406-68.
- (6) UNDERWOOD, E. (1931). *A Short History of French Painting*. London.
- (7) THOMSON, G. H. (1939). *The Factorial Analysis of Human Ability*. Univ. London Press.
- (8) THOMSON, G. H. (1936). "Some points of mathematical technique in the factorial analysis of human ability." *J. Exp. Psychol.* XXVII, 36-42.
- (9) THOMSON, G. H. (1940). *An Analysis of Performance Test Scores of a Representative Group of Scottish Schoolchildren*. Univ. London Press.

(Manuscript received 28 April 1944)

GRAPHOLOGICAL ANALYSIS AND PSYCHIATRY: AN EXPERIMENTAL STUDY

BY H. J. EYSENCK, *Psychological Department, Mill Hill Emergency Hospital*

I. *Introduction* (pp. 70-72). II. *The Experiment* (pp. 72-78): (1) *Method one: Questionnaire* (pp. 73-74); (2) *Method two: Intelligence test* (p. 75); (3) *Method three: Graphologist's matchings* (pp. 75-76); (4) *Method four: Psychiatrists' matchings* (pp. 76-77); (5) *Method five: Non-graphologists' matchings* (pp. 77-78). III. *Summary and conclusions* (pp. 78-79). *References* (pp. 79-80). *Appendix* (pp. 80-81).

I. INTRODUCTION

If the claims of graphologists to be able to diagnose temperamental and intellectual traits from a person's handwriting have any justification, it is clear that the scientific study of this discipline must be of outstanding importance to abnormal as well as to experimental psychology. The possibility of obtaining an analysis free from conscious or unconscious falsification on the part of the subject or the patient; the possibility of studying the pre-neurotic, pre-psychotic, or pre-experimental personality of subject or patient from objective records, such as early writings of his; the possibility, lastly, of following the progress of the cure, or of the effects of the experiment, in the image of the patient's or subject's handwriting—these are advantages enjoyed by no other method of personality study.

It is not surprising, therefore, that many psychologists and psychiatrists have advocated the use of graphological analysis in their respective fields; Schorsch⁽³⁶⁾, Alten⁽²⁾, Booth⁽⁴⁾, Gross and Bauer-Chlumberg⁽¹¹⁾, Lewinson⁽²¹⁾ and Theiss⁽⁴⁰⁾ are only a few of the more recent advocates in this respect. But the boundlessness of the graphologist's faith, the enormity of his claims, must make the cautious scientist hesitate. Unger's claim, for instance, to be able to set up graphological guides for the recognition of manic-depressive insanity, epilepsy, paranoia, paresis, schizophrenia, hysteria, neurasthenia, and alcoholism⁽⁴¹⁾ is liable to repel, rather than attract, the psychiatrist; similarly, Vertesi's claim to be able to diagnose liability to cancer from handwriting⁽⁴⁴⁾ will not impress the expert in cancer research, nor will Kroeber-Kenneth's view that accident-proneness can be diagnosed graphologically appeal to the industrial psychologist⁽¹⁹⁾.

Suggestive results have, however, been obtained by several workers who compared certain features in the handwritings of normal people with those found in the writings of schizophrenes and manic-depressives⁽²³⁾, or of dementia paralytica cases⁽³⁰⁾, or who compared the pre-psychotic writings of schizophrenes with their psychotic writings⁽³⁵⁾. Paskind & Brown⁽²⁹⁾, for instance, found various significant differences in height of letters, etc., between seventy-one deteriorated and forty non-deteriorated epileptics, and between these and various control groups.

Such group-differences, however, are of little help in individual analysis. Legrün⁽²⁰⁾ showed, for instance, that when the adolescent handwritings of a number of persons who later developed psychoses (mainly schizophrenia), and also of some seventy suicides, were analysed, none of the samples showed any abnormality, or even peculiarity, which was not also found in normal adolescents. Even an already developed psychosis, he found, did not always find expression in the handwriting.

While in these studies the validating criterion was usually a psychiatric rating, much work has been done in recent years by psychologists in which an attempt was made to validate graphological analysis against the results of experimental tests. The results have been very variable. Crider⁽⁸⁾ found an average correlation of 0.15 and 0.27 between ratings on sixteen traits of eighteen young adults by two graphologists, and by objective tests; Super⁽³⁹⁾ found no agreement between the judgements of a graphologist and the results of psychological tests on twenty-four college students; Harvey⁽¹⁸⁾ found that certain aspects of handwriting could be measured very reliably, but that their validity, when checked against various psycho-neurosis and other temperament scales, was very low (he used twenty-two measurable characteristics of the writings of his subjects). The validity could be improved, however, by pooling some of these characteristics, and by calculating multiple correlation coefficients.

Stackman⁽³⁷⁾ checked an objective record of "amount of activity" of 100 students against various features of their writing, without finding any significant differences. Cantril, Rand & Allport^(6, 7) found significant correlations between graphological analyses and scores on the Allport-Vernon 'Study of Values'. While this experiment may be thought to validate the 'type' theory held by Spranger, on whose views the 'Study of Values' was built, it is interesting to note that two other studies, those by Enke⁽⁹⁾ and by Schade⁽³⁴⁾, also claim to validate by means of checking graphological indices against objective tests the respective typological theories which are held by these writers, viz. those of Kretschmer and of Pfahler.

Intelligence tests and ratings of intelligence have been used by Hartge^(15, 16) and by Simon & Schönfeld⁽³⁸⁾ as a check on graphological analysis. Hartge's results show a certain amount of agreement between test results and graphological analysis. Simon & Schönfeld report a correlation of 0.33 from their study. They also gave a seventy-four item questionnaire to their subjects, to be filled in during a personal interview, and found that the graphologically determined traits were verified in 45 % of the cases, were doubtful in 21 %, and were wrong in 9 %. In 25 % of the cases the absence or presence of the trait could not be objectively diagnosed.

The ability of laymen to analyse handwritings has been studied by Theiss⁽⁴⁰⁾, Broom & Basinger⁽⁵⁾, and by Middleton^(24, 25, 26). Both Middleton and Broom & Basinger found that these judgements of both intellectual and temperamental functions had little reliability and hardly any validity; Theiss, on the other hand, claimed that the judgements of his subjects were considerably above chance.

Certain attempts have been made to improve the objectivity of various measurements and ratings in connexion with the analysis of handwriting. Wintz⁽⁴⁵⁾ has constructed an apparatus which allows of the measurement of pressure and speed in writing movements at the same time; Ruesch, Finesinger & Schwab⁽³²⁾ suggest methods for constructing an electromyogram of handwriting which may throw some light on problems of tension and relaxation; Lewinson & Zubin⁽²²⁾ have constructed elaborate scales which allow of the objective measurement of various aspects of a person's handwriting; Mira⁽²⁷⁾ has shown how certain handwriting movements can be dissected and studied in isolation and at will.

Of great interest is a study by Krueger & Zietz, summarized by Bachmann⁽³⁾, because it may throw some light on the vexed problem of just why in certain studies the results are so much more favourable to graphology than in others. The investigators laid before eighty people a 'universal' characterization, which was recognized by every one of their

subjects as appropriate for himself! A skilled writer, therefore, even without seeing the handwriting of a subject, should be able to write a character sketch which would be recognized by the subject as describing himself.

Also of some interest, from a methodological point of view, is a study by Ostermeyer & Sterzinger (28). These investigators studied changes in the writing of eleven subjects with increasing speed, with anger, and with dissimulation. This work would be well worth repeating on a larger sample.

The thirty papers mentioned in this brief review constitute what the writer considers the most important 10 % of the studies on graphology which have appeared since the Allport & Vernon *Studies in Expressive Movement* (1). Their conclusions, reached after a thorough review of the experimental evidence, have in the main been confirmed by later research. "Concerning the validity of graphological analyses of personality, there is evidence that controls which are stringent and at the same time permit reasonable exercise of the graphologist's talents, may eventually settle the problem to the scientist's satisfaction without necessarily overwhelming the graphologist with disaster" (1, p. 211).

Unfortunately, much of the evidence as reported in the various papers quoted cannot be regarded as more than suggestive. Too frequently the controls have not been sufficiently stringent, the number of handwritings used too small to give results free from serious sampling errors, and the criteria for validation themselves too much lacking in both reliability and validity to make comparisons fruitful. In the present investigation, an attempt was made to exclude errors from the sources mentioned as far as possible by (a) adopting the most stringent controls possible, (b) using a relatively large sample of handwritings, and (c) using criteria for validation which would be more acceptable than many in common use.

II. THE EXPERIMENT

The subjects of this investigation were fifty neurotic patients (male) at the Mill Hill Emergency Hospital. For each one of these patients, a case-history and the psychiatrist's diagnosis and personality analysis were available; these data constitute the primary validation criterion. Each man was also made to fill in a personality questionnaire in consultation with the experimenter; this questionnaire was constructed specially for the purpose in such a way that it would cover the main personality traits discovered by Guilford (12, 13, 14), and be intelligible to our patients. Also available for each man was his score on a standard non-verbal test of intelligence, the *Progressive Matrices* (31).

None of the writers knew the object of our study. They were asked to copy the personality questionnaire on standard, unruled white paper with a standard pen, and to write their answers behind the questions in a space previously marked off. When they had finished, the answers were detached, and the copy of the questionnaire in the patient's handwriting was then handed over to the graphologist. She was asked to fill in the answers to each of the questions in the questionnaire for each of the patients, to give intelligence ratings for the patients on a five-point scale, to write a character sketch of each, to be matched later by the psychiatrist against the man himself, and lastly to match the writings against the data supplied by the psychiatrist and by the patients' own answers to the questionnaire. (This last request was only made after all the other tasks had been completed, of course.) A more detailed description of these various procedures is given below, together with the results.

In addition to the four experiments mentioned, one further experiment was carried out in which ten psychiatrists and psychologists without any knowledge of graphology tried to match some of the writings against character sketches.

The graphologist, the results of whose work are reported here, was Dr O. Marum, formerly assistant in the Department of Psychology at the University of Bonn. During the first World War she worked in a mental hospital, and was thus better qualified psychologically and psychiatrically for her task than most graphologists.

(1) *Method One: Questionnaire.* In order to get self-assessments of the fifty patients on the major temperamental traits, they were asked to fill in the questionnaire given in the Appendix. Each question contrasts two different attitudes, ways of feeling, ways of acting, or general preferences; the subject is asked to write 'A' behind the question if the first alternative is the one that more nearly applies to him, and 'B' if the second alternative does. The patients were encouraged to elaborate on their replies by writing a fuller answer if they felt that simply writing 'A' or 'B' did not give a true picture of their feeling on the point; over 85 % availed themselves of this opportunity. They were also encouraged to ask the experimenter about any difficulties that arose in understanding the exact meaning of a phrase, or about the best way of answering the question.

The questionnaire was never done by more than two subjects at a time, so that individual attention could be given to each, to ensure as far as possible that they all understood the questions properly, and answered them correctly. The answers were then checked against the clinical case-notes of the patients, or gone over with the psychiatrist in charge of the case, and the general impression was that each patient had genuinely tried to give an honest picture of himself. Indeed, many patients elaborated greatly on their answers when talking to the experimenter, and emphasized their keenness to give the correct answers.

As pointed out above, the patients were made to copy the questionnaire in their own handwriting, apart from answering it; the answers were then detached, and the fifty questionnaires, in the various handwritings, handed over to the graphologist, who in turn attempted to *answer the questions for each patient*.¹ Question 18 ('Are you a bad sleeper...') the graphologist found impossible to answer, and it was therefore omitted from the following calculations. We then have altogether 1350 judgements by the graphologist, which can be checked against the judgements of the patients. In 152 cases, the graphologist was particularly sure of her judgement, and underlined the answers.

Scores were made up according to the following system: +2 points for each correct judgement underlined; +1 point for each correct judgement not underlined; 0 points for each judgement whose correctness was ambiguous (usually because the patient would not commit himself to saying either 'A' or 'B'); -1 point for each incorrect judgement not underlined; and -2 points for each incorrect judgement underlined. The number of judgements in these various groups is given in Table 1.

Taking into account only 'correct' and 'incorrect' judgements, regardless of underlining, and omitting ambiguous cases, we find that $62\% \pm 1.4\%$ S.E. of the graphologist's judgements are correct; this is significantly above chance beyond any reasonable

¹ It will be noticed that the handwriting specimens were obtained by *copying*, not by *free writing*. This was done to prevent judgements based on the writer's method of expression, choice of words, etc. from influencing the graphologist; in *free writing* it is impossible to control this factor adequately. It is possible that *copying* produces a different kind of writing, but this has never been satisfactorily proved.

doubt. Taking into account only those cases in which the graphologist was particularly certain of the accuracy of her judgement, the percentage of correct judgements rises to $68\% \pm 3.5\%$ s.e. Using plus and minus points, as shown in Table 1, the over-all percentage of correct judgements is $63\% \pm 1.0\%$ s.e. These values leave little doubt that something other than chance has played a part in the graphologist's success.

We can calculate the number of points scored by the graphologist on each of the twenty-seven questions, and on each of the fifty patients. The number of points scored by the graphologist for each of the questions is given with the question (cf. Appendix). On the whole, there are 352 points above chance, or an average for each question of 13. The

Table 1

No. of points	No. of judgements
+2	104
+1	669
0	100
-1	429
-2	48
Total	1350

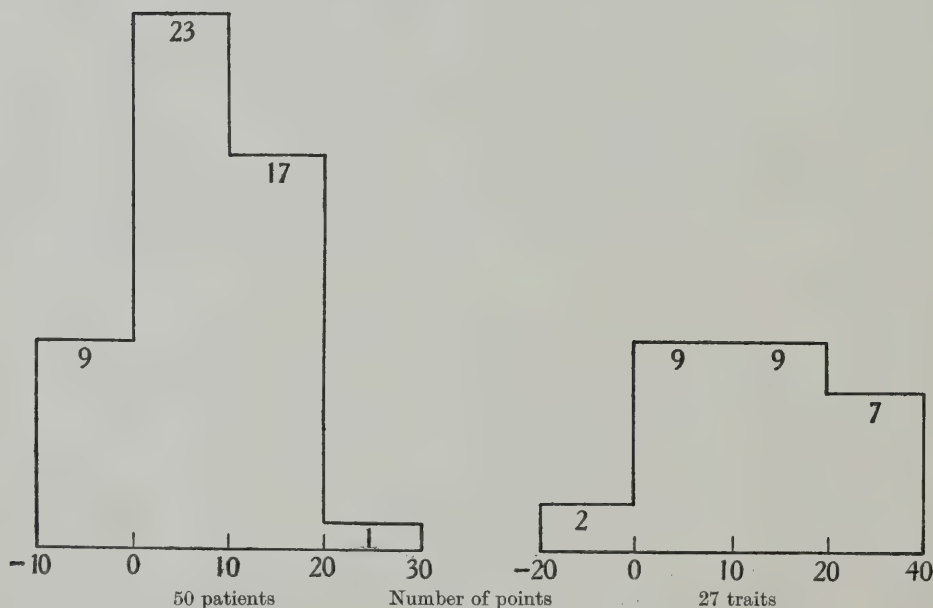


Fig. 1.

highest number of points for any one question is 39; the lowest -18 ('easily startled' and 'ponders over past' respectively).

The average number of points above chance scored by the graphologist on each patient is 7.04; the individual values vary from +26 to -10. Fig. 1 shows the distribution of scores for fifty persons and for twenty-seven questions respectively. It will be seen that both curves resemble a normal distribution curve, and that both extend into the region of minus scores. It seems reasonable to conclude (a) that persons vary with regard to the ease with which their handwriting may be diagnosed, and (b) that traits differ with regard to the ease with which their presence or absence may be detected from handwritings. Further evidence regarding these propositions will be presented later on.

(2) *Method Two: Intelligence test.* The Matrices Test of intelligence is given as a routine test to incoming patients, and the scores were available for checking against the graphologist's estimate of the patients' intelligence. (She was told the type of test her estimates would be validated against.) The test grades subjects into five groups, from Grade I (the very bright) to Grade V (the very dull). The graphologist followed this practice by giving each of the patients an intelligence Grade from I to V, based solely on his handwriting. Table 2 shows the average Matrix Test grades of patients classified by the graphologist in Grades I-V.

Table 2

Graphologist's grade	Average matrix grade	No. of cases
I	2.0	5
II	2.8	14
III	3.1	22
IV	2.4	7
V	3.0	2
Total		50

This table shows clearly that the graphologist's opinion does not coincide with the results of the test, except in the case of patients in Grade I. Of four patients who were in that grade, the graphologist picked out three; this is definitely statistically significant ($p < 0.01$). The two patients placed by her in Grade V, however, were actually both Grade III; similarly, of those she placed in Grade IV, none was actually graded as low as that.

It appears possible that what the graphologist attempted to diagnose was not 'ability to reason in abstract terms', or to 'deduce relations and correlates', but what Babcock has called 'mental efficiency'. I have discussed the problem of this distinction between 'level' and 'efficiency' of mental functioning elsewhere (10), in connexion with an analysis of the test scores of some 3000 neurotics, and believe that when graphologists talk about 'intelligence', it is the 'efficiency' aspect they deal with, rather than the 'level' aspect. (Actually, it might be more accurate to say that they combine these two aspects in proportions depending on the individual graphologist.) If this hypothesis be accepted, the lack of success of the graphologist receives an explanation which may serve to make later analyses more accurate and successful.

One point is of interest here which relates to the difference in 'ease of interpretation' of people's handwriting. In forty-one cases the graphologist was correct or one grade out in her estimation of the patients' intelligence; in nine cases she was out two or more grades. When we average the number of points the graphologist scored on these two groups of patients for the questionnaire test, the difference is very marked: the average number of points on the forty-one whose intelligence was guessed more or less correctly is 8.2, while the average number of points on the nine whose intelligence was misjudged completely is only 0.9! Thus those whose intelligence is misjudged completely are also those whose temperamental traits are judged hardly better than chance, and those whose intelligence is judged with a certain amount of success are also those whose temperamental traits are judged considerably better than chance.

(3) *Method Three: Graphologist's matchings.* For each of the fifty subjects of this study were available: (a) his clinical notes, containing a personality sketch as well as an account of his illness and neurotic history, (b) his answers on the temperamental questionnaire,

and (c) his scores on the intelligence test. The notes, answers, and scores belonging to each subject were put together and made identifiable by two code letters; names and every other means of identification were, of course, omitted. These various items of information regarding each patient will be referred to as his 'write-up'.

Ten chance samples of five write-ups each were then drawn, and the graphologist required to match, within each group, the five write-ups against the handwritings of the five patients concerned. Thus we have ten five-by-five matching problems. The results of this experiment are set out in Table 3. It will be seen that on the average 2.4 matchings were correct, as compared with a chance value of only 1. These results leave no doubt that the matchings are better than chance. The average coefficient of contingency¹ is 0.46; five out of the ten coefficients are more than three times as large as their standard errors, or more than five times as large as their probable errors.

Table 3

	No. of matchings correct	Coefficient of contingency
Group 1	3	0.71
Group 2	1	0.00
Group 3	1	0.00
Group 4	2	0.45
Group 5	5	0.90
Group 6	1	0.00
Group 7	3	0.71
Group 8	2	0.45
Group 9	3	0.71
Group 10	3	0.71
Average	2.4	0.46

(4) *Method Four: Psychiatrists' matchings.* The graphologist was asked to provide a short, one-page character sketch of each of the fifty subjects. These sketches were then to be matched by the psychiatrists in charge of the various cases against the personalities of the patients. As the number of patients who had been treated by the same psychiatrist varied considerably, it was impossible to keep the numbers of cases matched by each of them the same. The matchings were not $t : t$, as in the previous method, but $t : t'$; in other words, one sketch was introduced into the number of sketches the psychiatrist was required to match against his patients which did not apply to any of them. Thus the psychiatrist had to match, for instance, five sketches against four patients, throwing out the one sketch which did not apply to any of his patients. This procedure, of course, makes his task considerably more difficult.

Table 4

Psychiatrist	Type of matching	No. of correct matchings
Ch.	4 by 5	0
Re.	5 by 6	0
Gr.	4 by 5	1
Gr.	3 by 4	1
Gr.	4 by 5	2
Ep.	3 by 4	2
Ep.	4 by 5	2
La.	4 by 5	1
	Average	1.1

¹ The formulae for calculating this coefficient and its P.E. are given by Vernon (42, 43). The coefficient is analogous to a correlation coefficient, and roughly identical with it except towards the upper limits.

The results of this experiment are given in Table 4. It will be seen that the matchings were less successful than those carried out by the graphologist, and are not significantly better than chance. By chance we should expect an average of 1 correct matching, as compared with an experimental average of 1.1. The reason for this difference between the results of the two experiments, apart from the greater difficulty of the $t:t'$ procedure, is probably to be found in the fact that the cases in each particular doctor's ward tend to show much the same syndromatic picture; consequently the differences between them are probably less pronounced than they would be in a random sample, such as was used in method three. Such homogeneity in the sample tends to lower correlations considerably in this type of experiment.

(5) *Method Five: Non-graphologists' matchings.* Ten personality sketches were prepared by psychiatrists in charge of the patients in question, and divided into two groups of five each. The sketches in the two groups, called A and B, were given to ten psychiatrists and psychologists entirely (or almost entirely) ignorant of graphological principles, to be matched by them against the handwritings of the patients in the two groups. Thus we have altogether ten times two five-by-five matching problems. The results of the experiment are set out in Table 5.

Table 5

Judge	No. correct in	
	Set A	Set B
Han.	0	1
Mar.	1	2
Lew.	0	0
Hal.	0	0
Sen.	1	2
Jon.	1	0
Rav.	1	1
Cha.	1	2
Sti.	0	1
Sto.	0	0
Average	0.5	0.9

On the average, the number of correct matchings given by non-graphological judges is well below chance; in Set B, the average number of correct matchings is 0.9, while in Set A it is 0.5. By chance, we would expect an average of 1 correct matching in each set. These results should be compared with the average of 2.4 correct matchings by the graphologist, in a matching problem which for practical purposes was equal in difficulty to the one presented to our non-graphological judges. Out of a possible 50 correct matchings, the graphologist obtained 24; out of a possible 100, the non-graphologists obtained 14. It must be remembered, of course, that the graphologist had more material to aid her in her matchings than did the non-graphologists (sketches, questionnaires, and intelligence rating), but it is very doubtful if the additional material was really of any great help to the graphologist, who relied mainly on the personality sketch.

On the whole, the results reported in this paper seem to point to the conclusion that *it is possible for a skilled graphologist to diagnose personality traits from handwriting with better-than-chance success.* It seems also probable that the handwriting of some people presents less difficulty for such an analysis than does that of others, and that some traits are diagnosed with greater success than are others. This latter conclusion seems important because of the nature of the traits which graphologists sometimes claim to be able to

diagnose, although there may be little psychological evidence for the existence of a unitary trait of that kind; compare, for instance, Saudek's claim to be able to diagnose 'dishonesty' from handwriting (33), p. 376) with the results of Hartshorne & May's experimental studies of honesty and deceit (17). Clearly, if we accept the demonstration of these authors that dishonest behaviour is largely the outcome of a specific situation, and not of a generalized trait, then any attempt to diagnose such a generalized trait from handwriting or in any other way must be doomed to failure.

Altogether, it may be said that closer collaboration of graphologists and psychologists is desirable if we are to make any real advance in this particularly obscure and difficult field of personality assessment and 'expressive movement'.

III. SUMMARY AND CONCLUSIONS

The handwritings of fifty neurotic male Service patients at Mill Hill Emergency Hospital were analysed by a graphologist, and checked in a variety of ways against various personality indices. These included a personality analysis and diagnosis by the psychiatrist in charge of the case, a history of the neurotic illness of the patient, the patient's answers on a personality inventory administered personally, and the results of an intelligence test. The following results were arrived at, and the following conclusions drawn:

1. Out of 1350 judgements by the graphologist, concerning certain points of the patients' temperament and personality, $62\% \pm 1.4\%$ S.E. were correct, as checked against the patients' own answers. (Chance expectation = 50% .)

2. Out of 152 judgements by the graphologist on which she felt particularly confident, $68\% \pm 3.5\%$ S.E. were correct. (Chance expectation = 50% .)

3. The graphologist did not succeed in predicting the intelligence gradings of the patients, with the exception of those with very superior intelligence.

4. Persons tend to differ consistently with regard to the ease with which their handwriting can be diagnosed.

5. Traits differ with regard to the ease with which their presence or absence may be detected from handwritings.

6. The graphologist succeeded in matching character sketches of the patients, and other material, against their handwritings, the average coefficient of contingency being 0.46 for ten five-by-five matchings.

7. Psychiatrists did not match character sketches written by the graphologist on the basis of the patients' handwriting with the actual personalities of the patients with better-than-chance success.

8. Ten psychologists and psychiatrists with no, or very little, knowledge of graphology matched character sketches and handwritings with rather less-than-chance success.

9. Taken together, these results seem to show fairly conclusively that it is possible for a skilled graphologist to diagnose personality traits from handwriting with better-than-chance success.

10. Further progress in this field seems to depend largely on the close collaboration of graphologists and psychologists.

This investigation was carried out with the help of a Rockefeller grant. I am indebted to the Superintendent of Mill Hill Emergency Hospital for his permission to use the clinical material at the hospital. Lastly, it is a pleasure to record my gratitude to those

psychologists and psychiatrists whose kind assistance made the successful completion of this study possible.

REFERENCES

- (1) ALLPORT, G. W. & VERNON, P. E. (1933). *Studies in Expressive Movement*. New York: Macmillan.
- (2) ALTEN, E. H. (1939). "The psychology of handwriting, and its importance to the physician." *Med. Rec.* CL, 71-4.
- (3) BACHMANN, J. (1937). "Das graphologische Gutachten und seine Bewährung." *Z. Menschenkunde*, XIII, No. 3.
- (4) BOOTH, G. C. (1937). "The use of graphology in medicine." *J. Nerv. Ment. Dis.* LXXXVI, 674-9.
- (5) BROOM, B. H. & BASINGER, M. (1932). "On the determination of the intelligence of adults from samples of their penmanship." *J. Appl. Psychol.* XVI, 515-19.
- (6) CANTRIL, H., RAND, H. A. & ALLPORT, G. W. (1933). "The determination of personal interests by psychological and graphological methods." *Char. & Pers.* II, 134-43.
- (7) CANTRIL, H. & RAND, H. A. (1934). "An additional study of the determination of personal interests by psychological and graphological methods." *Char. & Pers.* III, 72-8.
- (8) CRIDER, B. (1941). "The reliability and validity of two graphologists." *J. Appl. Psychol.* XXV, 323-5.
- (9) ENKE, W. (1938). "Handschrift und Charakter im exakten Versuch." *Klin. Wschr.* XVII, 1624-7.
- (10) EYSENCK, H. J. (1943). "Neurosis and intelligence." *Lancet*, Sept. 18, p. 362.
- (11) GROSS, K. & BAUER-CHLUMBERG, M. (1937). "Handschrift und Geisteskrankheit." *Jb. Psychiat. Neurol.* LIV, 312-29.
- (12) GUILFORD, J. P. & R. B. (1936). "Personality factors S, E, and M, and their measurement." *J. Psychol.* II, 109-27.
- (13) GUILFORD, J. P. & R. B. (1939). "Personality factors D, R, T, and A." *J. Abnorm. (Soc.) Psychol.* XXXIV, 21-36.
- (14) GUILFORD, J. P. & R. B. (1939). "Personality factors N and GD." *J. Abnorm. (Soc.) Psychol.* XXXIV, 239-48.
- (15) HARTGE, M. (1933). "Graphologische Untersuchung der Intelligenz." *Z. angew. Psychol.* XLIV, 321-41.
- (16) HARTGE, M. (1938). "Graphologie in der Pädagogik und Berufsberatung." *Z. angew. Psychol.* LIV, 92-127.
- (17) HARTSHORNE, H. & MAY, M. (1928). *Studies in the Nature of Character. I. Studies in Deceit*. New York: Macmillan.
- (18) HARVEY, O. L. (1934). "The measurement of handwriting considered as a form of expressive movement." *Char. & Pers.* II, 310-21.
- (19) KROEBER-KENNETH, L. (1938). "Unfallneigung und Handschrift." *Z. Menschenkunde*, XIV, 17-32.
- (20) LEGRÜN, A. (1939). "Schriften Geisteskranker aus ihrer Jugendzeit." *Z. Pädag. Psychol.* XL, 85-91.
- (21) LEWINSON, T. S. (1940). "Dynamic disturbances in the handwriting of psychotics; with reference to schizophrenic, paranoid, and manic-depressive psychoses." *Amer. J. Psychiat.* XCVII, 102-35.
- (22) LEWINSON, T. S. & ZUBIN, J. (1942). *Handwriting Analysis*. New York: King's Crown Press.
- (23) MANDOWSKY, A. (1932). "Vergleichende psychologische Untersuchungen über die Handschrift." *Zbl. Psychotherap.* V, 715-33.
- (24) MIDDLETON, W. C. (1939). "The ability of untrained subjects to judge dominance from handwriting samples." *Psychol. Rec.* III, 227-38.
- (25) MIDDLETON, W. C. (1941). "The ability of untrained subjects to judge intelligence and age from handwriting samples." *J. Appl. Psychol.* XXV, 331-40.
- (26) MIDDLETON, W. C. (1941). "The ability of untrained subjects to judge neuroticism, self-confidence, and sociability from handwriting samples." *Char. & Pers.* IX, 227-34.
- (27) MIRA, E. (1939). "Myokinetic psychodiagnosis: a new technique of exploring the conative trends of personality." *Proc. Roy. Soc. Med.* XXXIII, 173-94.
- (28) OSTERMEYER, G. & STERZINGER, O. (1937). "Graphologische Untersuchungen." *Z. angew. Psychol.* LII, 1-23.
- (29) PASKIND, H. A. & BROWN, M. (1940). "Constitutional differences between deteriorated and non-deteriorated patients with epilepsy. IV. The handwriting." *Arch. Neurol. Psychiat.* XLIV, 507-15.

- (30) QUINAN, C. (1931). "The time and lineage factors of handwriting." *Arch. Neurol. Psychiat.* xxvi, 333-45.
- (31) RAVEN, J. C. (1941). "Standardization of progressive matrices, 1938." *Brit. J. Med. Psychol.* xix, 137-50.
- (32) RUESCH, J., FINESINGER, J. E. & SCHWAB, R. S. (1940). "The electromyogram of handwriting." *Psychosom. Med.* II, 411-37.
- (33) SAUDEK, R. (1928). *Experiments with Handwriting*. London: Allen and Unwin.
- (34) SCHADE, W. (1939). "Handschrift und Erbcharakter." *Z. angew. Psychol.* lvii, 303-81.
- (35) SCHORSCH, G. (1938). "Die Ausdruckskunde in der psychiatrischen Diagnostik." *Zbl. ges. Neurol. Psychiat.* lxxxvii, 705.
- (36) SCHORSCH, G. (1939). "Die prämorbid Persönlichkeit bei Schizophrenen." *Z. ges. Neurol. Psychiat.* clxvii, 154-7.
- (37) STACKMAN, H. A. (1934). "Handwriting and extra-curricular activity." *J. Appl. Psychol.* xviii, 819-25.
- (38) SIMON, W. & SCHÖNFELD, W. (1935). "Vergleich der Intelligenzschätzung mit graphologischer und psychotechnischer Methode." *C.R. 8. Conf. int. Psychotech.* Prague, 700-1.
- (39) SUPER, D. E. (1941). "A comparison of the diagnoses of a graphologist with the results of psychological tests." *J. Consult. Psychol.* v, 127-33.
- (40) THEISS, H. (1931). "Experimentelle Untersuchungen über die Erfassung des handschriftlichen Ausdrucks durch Laien." *Psychol. Forsch.* xv, 276-358.
- (41) UNGER, H. (1935). "Geisteskrankheit und Handschrift." *Z. ges. Neurol. Psychiat.* clii, 569.
- (42) VERNON, P. E. (1936). "An evaluation of the matching method." *J. Educ. Psychol.* xxvii, 1-12.
- (43) VERNON, P. E. (1936). "The matching method applied to investigations of personality." *Psychol. Bull.* iii, 149-77.
- (44) VERESI, E. (1939). *Handschrift und Eigenart der Krebsgefährdeten*. Budapest: Tisza.
- (45) WINTZ, J. (1938). "Druck- und Geschwindigkeitsverlauf von ganzheitlichen Schreibbewegungsweisen." *Neue Psychol. Stud.* xi, 5-55.

APPENDIX

Points

- 8 1. Are you more interested in sports or in intellectual things?
- 6 2. Do you particularly dislike being "bossed", or do you generally do as you are told, without feeling rebellious?
- 1 3. Do you usually "kick up hell" when you are not getting a square deal, or do you usually just shrug your shoulders and say: "Oh, it doesn't matter"?
- 4 4. Do you enjoy being in a position where you can give orders, or do you dislike having such responsibility?
- 10 5. Do you consider yourself a happy-go-lucky individual, or are you inclined to worry over possible misfortunes?
- 0 6. Are you inclined to stop and think things over before acting, or do you often act on impulse?
- 18 7. Are you inclined to ponder over your past, or do you rather "live for the day"?
- 17 8. Are you rather over-conscientious in your work, or do you tend to be a bit "slap-dash"?
- 12 9. Are you inclined to try and watch your own mind at work and analyse your motives for doing things, or do you consider that sort of thing rather morbid?
- 6 10. Do you enjoy thinking out complicated problems, such as crossword puzzles, etc., or do you consider such things "highbrow"?
- 17 11. Do you adapt yourself easily to new conditions, or do you take quite some time to settle down?
- 35 12. Do you get rattled easily in exciting situations (examinations or manœuvres, for instance), or do you usually remain "calm, cool and collected"?
- 19 13. Do you often have feelings of inferiority, or do you usually feel self-confident?
- 26 14. Are you often troubled by feelings of self-consciousness, or do you go along happily without bothering much about what others will think of you?
- 2 15. When you are working, are you easily distracted by things that are happening around you, or do you usually concentrate whole-heartedly on your work?

- 8 16. When you are sitting or lying down, can you usually relax easily, or do you remain tense for a long time?
- 39 17. When something unexpected happens, are you easily startled, or do you remain quite unaffected?
- 18. Do you often sleep badly, or would you call yourself a fairly sound sleeper?
- 1 19. Would you consider yourself quick and agile in your actions, or are you rather “slow but sure”?
- 20 20. Do you often rush from one activity to another without pausing for a rest, or do you prefer having a bit of a rest before starting out on something new?
- 10 21. Do you like to talk a lot, or do you rather let the others do the talking?
- 9 22. Do you like meeting new people, or does the prospect of doing so rather upset you?
- 14 23. When you and your friends are doing something, do you usually take the lead, or do you rather fall in with other people’s wishes?
- 31 24. Do you have frequent ups and downs in mood, or do you go along “on an even keel”?
- 33 25. Are your feelings easily hurt, or do you tend to remain unruffled in general?
- 10 26. Do you tend to express such emotions as delight, sorrow, anger, etc., readily, or do you feel that it would never do to exhibit your emotions publicly?
- 14 27. When you are out with friends, do you enter into the fun whole-heartedly, or do you tend to be rather quiet and withdrawn?
- 20 28. Would you like to go through thrilling experiences and adventures yourself, or do you prefer reading about them, or seeing them at the pictures?

(Manuscript received 16 May 1944)

THE ENVIRONMENT AND THE CHILD

BY H. BANISTER AND M. RAVDEN

From the Child Guidance Clinic, Cambridge

I. *Preliminary* (p. 82). II. *Procedure* (p. 82). III. *Results* (pp. 82-83). IV. *Discussion* (pp. 83-84). V. *Further discussion* (pp. 84-86). VI. *Summary* (pp. 86-87). *References* (p. 87).

I. PRELIMINARY

In a previous paper⁽¹⁾ the home environment of 123 children who had been referred to the Cambridge Child Guidance Clinic was discussed. It seemed desirable to obtain information regarding the type of home and the environment of a control group of children who had not been referred to the clinic, and who were considered well-adjusted. This paper contains the results of this inquiry.

II. PROCEDURE

Five of the Cambridge Elementary Schools, typical of the schools from which the problem children came, were visited, and the registers were gone through with the head teachers. The name of the first child in a particular class who was considered satisfactory was noted and then that of the second and so on till the quota from that school had been obtained. The only provisos were that the distribution of the children over the whole school-age group and the proportions of boys and girls and of residents and evacuees, should be approximately those of the problem children.

Having in this way obtained the names of a number of children whom the head teachers considered satisfactory, the parents were approached with the request that they might be visited. In spite of the difficulty, for many of the parents were working all day, 100 out of 107 parents welcomed the suggestion. As a rule but one parent, and that usually the mother, was seen. In two cases the father only was interviewed and in fourteen instances both parents were seen together.

At the interview, which lasted from 1 to 2½ hours, the talk was directed so as to obtain a general picture of the home background, and of the parents' relationships to each other and the children. Conversation centred on their practices and their views concerning the discipline of their own and of other peoples' children; the role of the father and of the mother, and how far the children helped or should help with household duties; the children's activities; their amusements and friends and how far the parents attempted to control these; the conditions, if any, under which the children were permitted to go out alone; whether or not the children had ever presented any special problems of behaviour or sickness, and so on.

III. RESULTS

Since many of the problem children were well-behaved at school and presented difficulty only to their parents, it was expected that some of the children considered satisfactory by the teachers would not be thought satisfactory at home. This expectation was

justified; seven of the hundred children were presenting serious symptoms according to their parents, so that data concerning them have been excluded.

One other matter of interest, evacuees. 38 % of the problem children were evacuees, and it was suggested in the former paper⁽¹⁾ that the evacuee children still left behind in Cambridge might, in large part, be those whom the parents did not wish to have back because they were difficult. Support is given to this conjecture by the difficulty experienced in obtaining the quota of evacuees. One school had only twenty-seven evacuees left in February 1943 out of 250 at the outbreak of the war, and most of these twenty-seven had given trouble.

The homes of the ninety-three normal children, classified under the following headings: Accord (both parents living together in harmony); Discord (both parents living together, but with evidence of discord); One parent dead; Other broken homes (home broken by divorce, desertion, etc.), were:

	Accord	Discord	One parent dead	Other broken homes	Total
Boys	56	1	6	2	65
Girls	23	1	1	3	28
Total	79	2	7	5	93

IV. DISCUSSION

For the problem children the distribution of the 112 homes was:

Accord	Discord	One parent dead	Other broken homes	Total
52	15	15	30	112

Examination by the χ^2 method shows that the chance of the two distributions being drawn from the same population is less than 1 in 1,000,000. The differences between the two distributions cannot, therefore, be due to chance.

The criticism immediately arises that the classification of the homes of the satisfactory children may have been biased, that there was a tendency on the part of the assessor to consider the home to be in the 'accord' class when a more careful and prolonged inquiry would have shown that there was discord between the parents, and that in any case the parent or parents would certainly hide any conjugal differences from the interviewer.

Assuming this criticism to be valid, and taking only the truly objective criteria of whether both the parents are alive and living together, or whether one parent is dead, or whether the home is a broken one from other cause, we get the following distributions:

	Both parents alive and living together	One parent dead	Other broken homes	Total
Normal children	81	7	5	93
Problem children	67	15	30	112
Total	148	22	35	205

Examination of these data by the χ^2 method shows that the chance of the two distributions being random selections from the same population is less than 1 in 33,000 and that the contribution of the homes where one parent is dead to this result is very small.

The further criticism may be raised that this result is due to the inclusion in the problem

group of the delinquent children whose poor home environment has often been shown to be significant. Excluding the delinquent children, the following data are obtained:

	Both parents alive and living together	One parent dead	Other broken homes	Total
Normal children	81	7	5	93
Problem children	62	9	14	85
Total	143	16	19	178

The chance of these two distributions being random selections from the same population is less than 1 in 66, which is usually considered to be significant.

There can be little doubt therefore that the broken home is one of the factors associated with the production of the problem child, both delinquent and non-delinquent.

This fact has long been recognized in the case of the delinquent child, and has, in this connexion, been variously interpreted. Mullins ((7), p. 125) and others seem to consider that the broken home is the causative factor in the production of delinquency, which Healy & Bronner ((3), p. 208) ascribe largely to the resulting 'lack of normal parental supervision, guidance, companionship and control'; others, e.g. Hirsch ((4), Chap. v), suggest that the root cause is 'the poor constitutional make-up of the parent or parents' which is transmitted to the children, or to some of the children—since not all the children of a given family are delinquent—who are thus predisposed to delinquency. This would seem to place the greatest emphasis on a genetic factor which is so clearly brought out by Lange's study of twins(6). Yet the possibility is that the broken home, other than the accidental one caused by death, is often the result of instability in one or both parents, and that the instability produces a home environment conducive to delinquency. Though this does not rule out any inherited factor or predisposition, it may well be that in many cases parental instability acting in this way is the more important. Support is given to this supposition by the fact that broken homes are, as is shown above, associated with other evidences of instability than delinquent behaviour. This suggestion is in line with the results of Huschka's(5) inquiry into the psychopathological disorders of the mothers of problem children. An examination of the mothers of 488 problem children showed that 203 (41 %) were suffering from psychopathological conditions. In each case where the mothers had been examined or treated in the psychiatric department, the condition was an important factor in the problem of the child. 'The sick mother', she states ((5), p. 83), 'is one of the crucial factors operating daily to lay the foundations in her child for the very illnesses we seek to prevent.'

There would seem to be little doubt but that a very large proportion of the parents of problem children are in need of treatment, and that they often need the treatment more urgently than their children.

V. FURTHER DISCUSSION

In the previous paper(1) tentative suggestions as to the causes of their difficulties were made as a result of the analysis of the data regarding the problem children. Of the normal children some come from broken homes, some have only one parent alive, and two are from homes in which there was some evidence of discord. Is it possible to discover why these children are not finding life difficult? And is any information to be obtained from the 'accord' home?

(a) *The discordant home.* Only two normal children came from homes in which there was evidence of discord. One of these, a boy, showed evidence of minor difficulty, scarcely sufficient to necessitate his being sent to the clinic; the other child, a girl of 5, appeared to be well-adjusted, but an elder sister had been given to stealing. From these two cases no valid deductions can be drawn.

(b) *The home in which one parent is dead.* There were seven such homes, and one of the children, a boy, was said to be rather aggressive, but the family doctor attributed this to physical causes. The other six children were well-adjusted.

It is perhaps significant that none of these children was an only child: they all had older siblings, and in five instances the older children had to a certain extent taken over the functions of the dead parent. In every case the surviving parent was kindly and affectionate. Two of the parents were exceptionally understanding. They seemed very intelligent, took a keen interest in their children's welfare, and were determined that the children should not suffer because of their loss. In another case, an exceptionally competent grandmother, who had herself brought up eleven children, had filled the gap in the family circle most effectively, compensating fully for the loss of the children's mother. It was particularly noticeable that in none of these cases did the surviving parent endeavour to 'absorb' their children or prevent them from living their own lives fully. The children were encouraged to be independent and to engage fully in numerous outside activities, e.g. Scouts and Guides; their hobbies, e.g. making toys and model aeroplanes, were fostered and so on.

The key-note to success lay in wise handling: affection but not over-protection, kindly discipline combined with freedom, the fostering of various activities and no morbid concentration on the loss the family had suffered—a cause of difficulty which was found repeatedly in the homes of the children who were referred to the clinic and who had lost a parent. A further contribution to success would seem to be the replacement of the lost parent by the grandmother or by an older sibling, which facilitated the wise handling.

(c) *Homes broken by separation, divorce, etc.* Five normal children came from these homes. Two, evacuees, had been very difficult, had attended Child Guidance Clinics in the past, and gave much trouble when they came to Cambridge. They were put into good billets and are said to have changed completely since their evacuation. Two others, also evacuees, had given no trouble at any time; the one remaining child lived with his father, apparently a very sensible man, his grandmother, a kindly and understanding person, and his three brothers and sister in what appeared to be a very happy *ménage*.

In these cases it would seem that the crucial factor was the making of a substitute family to replace the broken one. The two evacuees who had been difficult before coming to Cambridge had been in Homes prior to their evacuation. It seems possible, as some suggest, that life in a Home, where shelter, food and clothing are provided, is no adequate substitute for a happy home life with the child's own parents. But this is a question on which sufficient evidence is not presented, in this and the preceding paper, to justify any conclusion.

(d) *The 'accord' home.* The outstanding features of this class of home call for little remark. In the best ones the 'atmosphere' was happiness, all the members of the family obviously pulling together. Parents often remarked on the fact that, if they had differences, they avoided letting the children know of them and settled them in private. They insisted on the importance of presenting what they called a 'united front' both between

themselves and towards their children, adhering rigidly to any decision they had taken with regard to the children, never giving vent to threat of punishment unless they were prepared to carry out the threat. They allowed the children to make their own decisions as far as possible in matters which concerned them, allowed them complete control over their pocket money, and did not fuss over them nor coddle them. Many of the mothers expressed the view that they had learned by experience and had allowed the younger children greater freedom than the eldest and had noticed what a difference this had made to their development. In many cases, realizing the danger of jealousy, parents took active steps to avoid giving any cause for that emotion to be aroused.

(e) *The only child.* It is interesting that the same proportion of only children is found in this group as was found among the problem children referred to the clinic, viz. 22 %. This bears out the well-recognized fact that the only child is not necessarily handicapped because he has no brothers or sisters, though it may be that further inquiry, with much larger numbers, would show that the only child was at a disadvantage if one parent were dead (cf. (b) above).

(f) *Children's activities.* There was one notable difference between these children and the problem children referred to the clinic—their activities and amusements. The problem children seemed usually to be at a loose end, not knowing what to do with their spare time: they 'play about', or, 'there's nothing to do', they 'don't like going to Scouts', they are 'bored', they 'like to go into the town', they 'ask soldiers for money to go to the cinema'; whereas these other children seem to be always occupied with activities, hobbies and games which absorb their energies fully and happily. It is not without significance that many of the problem children seem to find their chief amusement and relaxation at the cinema, which plays a comparatively small part in the lives of the normal children. Only about 40 % of the latter visit the cinema at all regularly, ranging from once a week to once a month, and in only two cases did the children go more than once a week. This finding bears out the results of previous inquiries (cf. Hardy (2)). It is not suggested that the cinema exerts a bad influence on the children, but that frequent attendance is probably indicative of some difficulty in the child's life.

VI. SUMMARY

1. Information regarding the homes of 93 normal elementary school children is presented and compared with the home conditions of 112 children who had been referred to the Cambridge Child Guidance Clinic.

2. The distribution of the homes in the two cases is shown, statistically, to be significantly different. It is not a chance factor that many problem children come from broken homes. It is suggested that this association may to a considerable extent be due to the psychological effects of instability in the parents.

3. Some satisfactorily adjusted children come from broken homes of various kinds. These homes are discussed. Satisfactory adjustment seems to depend, largely, on the provision of a close psychological substitute for a normal home.

4. The attitudes of parents in a normal home are outlined.

5. The proportion of 'only' children is the same in both groups, indicating that the 'only' child is not handicapped thereby.

6. The well-adjusted children have much more strongly developed interests, hobbies

and social activities than the problem children. In particular there is a much less frequent attendance at cinemas.

REFERENCES

- (1) BANISTER, H. & RAVDEN, M. (1944). "The problem child and his environment." *Brit. J. Psychol.* xxxiv, 60-5.
- (2) HARDY, M. C. (1934). "Out-of-school activities of well-adjusted and poorly adjusted elementary school children." *Psychol. Bull.* xxxi, 750.
- (3) HEALY, W. & BRONNER, A. F. (1926). *Delinquents and Criminals*. New York: Macmillan.
- (4) HIRSCH, N. D. M. (1937). *Dynamic Causes of Juvenile Crime*. Cambridge, Mass.: Sci.-Art Publishers.
- (5) HUSCHKA, M. (1941). "Psychopathological disorders in the mother." *J. Nerv. Ment. Dis.* xciv, 76-83.
- (6) LANGE, J. (1931). *Crime as Destiny*. London: George Allen and Unwin, Ltd.
- (7) MULLINS, C. (1943). *Crime and Psychology*. London: Methuen and Co. Ltd.

(Manuscript received 23 May 1944)

PUBLICATIONS RECENTLY RECEIVED

Figural after-effects: an investigation of visual processes. By W. KÖHLER and H. WALLACH.
Proc. Amer. Philosoph. Soc., 1944, LXXXVIII, 269-357.

The authors start from Gibson's discoveries—that prolonged inspection of slightly curved or tilted lines causes other straight or vertical lines, subsequently observed, to appear curved or tilted in the opposite direction. They proceed to a general investigation of visual after-effects. The main phenomenon which they report consists of a tendency for inspection of a black line to cause an apparent weakening and a measurable displacement away from the first line, in any subsequent line which is looked at. This effect is greatest when the second or 'Test' line is at a distance of about 0.75 in. from the first or 'Inspection' line, when looked at from 3 yd. (i.e. a separation of about 25 min. of arc). They consider that this is the basic effect both in Gibson's situations and, as a simultaneous rather than an after-effect, in Ponzo's and other visual illusions, and advance an explanation more general than that of Gibson. Gibson had accounted for his results by postulating a tendency for lines, on prolonged inspection, to approximate to certain 'norms' such as straightness and verticality; consequently any other lines subsequently regarded should appear distorted in the opposite sense by a sort of contrast effect. Köhler & Wallach show that this explanation will not suffice; for instance, the tilting effect occurs whatever angle the original lines may have been presented at. They postulate instead that the displacement of the second line by the first is the basic effect, and that it is due to a state of 'satiation' or 'depression' of the visual field produced by the first line, which will affect subsequent test figures provided they have not themselves configurations which conflict with such effects.

They suggest a further explanation of this 'satiation' or 'depression' in terms of 'molar physiology'. They postulate that the most stimulated regions of the visual cortex—where the bright part of the image falls—has its electrotonic resistance raised and consequently that there may be local differences in potential-distribution, causing transverse currents to flow. They quote Gerard & Libet's discovery of electrical waves which spread over the whole of a frog's cortex, even when all the neurones in their path have been cleanly transected, in support of their postulate. Such currents would be 'relationally determined', i.e. produced by *differences* in brightness of parts of the image and not by the *absolute value* of that brightness; and it is suggested that their interaction (its exact nature being intentionally left vague) may cause the displacement effects observed.

The paper is an important contribution to the theory of visual after-effects, illusions, and visual perception in general. The displacement which they have measured seems indeed to be the common factor of a great number of visual after-effects and illusions and a more general one than the "approximation to norms" and subsequent "contrast effects" of Gibson. Further, their emphasis on "relational determination" and transverse effects in the nervous system as essential parts of any theory of perception is commendable. Finally, their physiological theory, though very bold and admittedly tentative, does not seem inconsistent with any known physiological facts, and should be capable of some experimental test as electrophysiological methods improve. Their theory seems to me, however, to be somewhat warped by a sense of conflict between molar and atomic physiology which should not exist. Their desire to maintain this conflict, and to emphasize the merits of the 'molar' school, leads them both to undervalue the work of most electrophysiologists and also to think that molar action can accomplish more than in fact it can. They themselves seem to be aware that some of the Gestalt views can no longer be held in an extreme form; for instance, they admit that the 'nativistic' versus 'empiristic' controversy becomes "seriously misleading" since "visual objects may originally be quite different from what they are at a later stage" though "the change will not be attributable to anything like learning in the usual sense", and again that "*microscopically*, the visual brain remains an extraordinarily heterogeneous medium". They also quote results on, for instance, "relational determination" in the nervous system which were obtained by physiological work on isolated nerve fibres and other preparations. Yet after these admissions they plunge straight into the concept of molar currents in the cortex, justifying themselves in disregarding neuronal activity henceforward by saying that, on it, "impulses in parallel chains of neurones would have to be transversely interrelated across corresponding stretches of tissue" and that this "hypothesis, far from saving the customary interpretation of brain function, actually sacrifices the main principle of this interpretation". This last statement is highly questionable; it was physiologists working on isolated preparations who discovered such phenomena as synchronous activity of the fibres in whole nerves, excitation of one crustacean nerve-fibre by another lying next to it, convergence of neurones on motor-neurone pools, functional convergence of retinal receptors on optic nerve-fibres, and so forth. Besides being thus unfair to the labours of physiologists who have never regarded isolated preparations as equivalent to the whole organism but have very definitely found in them a useful means of attacking the problems of the action, integration, and coordination of the various organic elements, their bias seems to me to warp their theory in two ways.

First, by leaving a gap between the function of single neurones and the behaviour of molar currents in, as it were, a homogenous medium, it ignores the immense possibilities afforded by the interaction of large numbers of units. For instance, it is ordinarily thought to be a delicately balanced interplay of excitation

and inhibition which produces a co-ordinated spinal reflex; and again, in vision, it is possible that an interaction between neighbouring receptors in the retinal mosaic, such that a bipolar neurone is excited by the *difference* of excitation of the two receptors, would provide a possible contour-responsive mechanism. (This may be supported on the ground that this would provide a spatial method of 'differentiation', in the mathematical sense, and therefore be equivalent to the characteristic, which sense organs possess, of differentiating with respect to time, and so detecting sudden changes of illumination, as has been emphasised by Granit and by the present writer, and is of course 'relational determination' in time instead of in space.)

Secondly, the excessively 'molar' bias seems to lead Köhler & Wallach to disregard the principle of 'levels' which is so well-established and so valuable in the nervous system. For instance, some retinal adaptation or fatigue and after-effects do undoubtedly occur, and to disregard it is to invite criticism. Köhler & Wallach say that "if such a banal explanation can be given for one of the three symptoms one might be tempted to interpret the other two as secondary consequences of the first. They too would then become unimportant... however, it is not plausible even in the present instance." Not only is such emotional language rather unbecoming; it invites the retort that, although the displacement effect, as they point out, cannot be retinal since it holds when one eye looks at the 'Inspection' object and the other at the 'Test' one, the first effect (reduced contrast) is largely peculiar to one eye, as ordinary adaptation and after-images are known to be. Why deny that the nervous system involves a number of levels, of which the lower ones fulfil the simpler functions? The same criticism applies to Köhler & Wallach's attempt to explain away some results of Zangwill and the present writer, in which the rise in differential brightness threshold within a closed figure was shown not to be dependent on 'closedness' of the figure, but only on the presence of lines. Why not consider the possibility that brightness-discrimination may be peripherally determined, and seek better fields for testing the other, undoubtedly central, effects?

In short, the experiments and theory form an important contribution to the understanding of visual perception, synthesizing a number of psychological and physiological facts, but are, it seems to me, warped by the artificial gap which the authors fix between atomic and molar methods—between the study of the elements of the organism and their integrated action in the whole—so that the functions probably performed by neuronal and synaptic interaction, and by segregation of function into levels, are disregarded.

K. J. W. CRAIK

Persistence and Change in Personality Patterns. By KATHERINE ELLIOTT ROBERTS and VIRGINIA VAN DYNE FLEMING. Washington, D.C.: Society for Research in Child Development, National Research Council. 1943. Pp. vii+206. \$1.50.

In this monograph, twenty-five life histories of women obtained through interviews given by an Advisory Service on personal problems are analysed to determine how far personality traits and characteristic attitudes persist throughout life, and how far they are modified. Three of the life histories are given in full, and others described more briefly. A long list of traits and attitudes was drawn up and checked for each case by independent investigators, as they seemed to occur in the life history during three periods of life, pre-college, college and post-college. The women themselves checked a similar list, but the correlations for persistence of traits between the two sets of lists was low. It is not clear, therefore, that such conclusions as were reached could not have been equally well derived directly from studying the life history without resorting to this elaborate and only pseudo-objective procedure. As might be expected, some traits were persistent in some cases and others in others; some were comparatively persistent in all cases, others hardly appeared at all. There seems to have been rather more persistence on the whole than fluctuation. Little attempt was made, however, to show why certain traits were relatively modifiable and in what circumstances, beyond the one or two rather obvious facts such as that a happy life in childhood is likely to lead to stability and happiness in later life. The fundamental defect of this method of tackling the problem of the persistence of traits is, of course, that all evidence of behaviour characteristics and important modifying influences during the earliest years of life is necessarily lacking.

Branch Street. By MARIE PANETH. London: George Allen and Unwin, Ltd. 1944. Pp. 128. 6s. net.

This book describes the attempt made by Mrs Paneth, an Austrian refugee, to run a club or 'centre' for children living in a London slum area. This area Mrs Paneth calls 'Branch Street', and it is clearly one of the plague spots of London (it is *not* situated in the East End). The children were poverty stricken in the extreme, because their parents belonged to Charles Booth's Class A—the waster and semi-criminal; their behaviour was antagonistic, aggressive, destructive and lewd to a degree which Mrs Paneth says she had encountered among no other London slum children. She attributes it to the fact that once past babyhood they received no care and little affection; they had to fight and scrounge for everything they needed; their hand was against every man's and every man's against them. It was true that they were exceedingly kind to their baby brothers and sisters; but the care of these younger children imposed an

added strain on them. Mrs Paneth realized that any attempt to impose normal control and discipline led merely to further antagonism. And therefore when they came to the 'centre'—held first in an air-raid shelter and later in a house in 'Branch Street'—they were allowed to scream, fight and destroy at will. It did appear after a time that her calmness and friendliness in this appalling *mêlée* began to produce some effect; the children began to tire of their lewdness and destructiveness, and some of them, the girls especially, became interested in painting, sewing, acting and the like. The boys, too, consented to learn boxing. But alas, from lack of support from helpers and from the authorities the experiment was not continued in the manner and the spirit with which Mrs Paneth had begun it. Whether indeed the children would have become sufficiently socialized, in spite of their appalling home background, to accept the normal disciplinary methods which she intended at length to apply, Mrs Paneth herself is uncertain. Perhaps only a thorough break-up of the neighbourhood would have been effective—the 'worst' children sent to good foster homes and given psychotherapeutic treatment.

The book does nevertheless give an indication of what can be done by such treatment, given sufficient courage, endurance and charity. It is a very moving document; and a more horrifying indictment of our so-called civilization than was even *Our Towns*.

Crime and Psychology. By CLAUD MULLINS. With an Introduction by Dr Edward Glover. London: Methuen and Co., Ltd. 1943. Pp. xvi+234. 8s. 6d. net.

This is not an ordinary book on the psychology of crime: it is a book with a purpose, no less a one than the reform of almost the whole of our present system of dealing with crime. When one realizes that even now the medical practitioner still may have no appreciation of the value of psychology and will advise parents on no account to send their children to a Child Guidance Clinic, one appreciates the immensity of the task which Mr Mullins has undertaken, when he wishes to reform the outlook of a still more conservative profession, and one to which the applications of psychology are not so obvious. Among the many changes he suggests, he wishes the study of psychology to form part of the training of the police force and of the legal profession, not that the policeman and the lawyer may become expert psychologists, but that they may have some appreciation of the springs and of the vagaries of human behaviour. He would do away with trial by jury. He places all the emphasis on the reform of the criminal and not on his punishment, though that may be advisable in certain circumstances. He advocates a delay in general between finding a verdict of guilty and the pronouncement of sentence in order that a full clinical examination of the prisoner may be carried out; and, acknowledging the difficulties, he seeks the co-operation of all, and particularly of the psychologists, to help in the reform. His well-documented argument is very convincing and is well illustrated by cases which have come before him mostly at the South-Western Court.

Crime and Psychology is an important book which should be widely read. The need for the reforms is evident to most of those who have ears to hear and eyes to see and who have had experience of the ways of dealing with delinquents, young and old.

H. B.

Psycho-Analysis and Crime. By S. H. FOULKES. With a Preface by CYRIL BURT. Department of Criminal Science, Faculty of Law, University of Cambridge: English Studies in Criminal Science, Pamphlet Series. Published by the Canadian Bar Association. 1944. Pp. 43.

The author begins with a brief yet clear outline of present-day psycho-analytic doctrines and methods, and goes on to discuss delinquency from this standpoint, drawing on his own researches, as well as on those of others, to illustrate the various points. He then considers law, delinquency and their psychology as social phenomena.

The pamphlet is a real contribution to the study of the psychology of crime.

PROCEEDINGS OF THE BRITISH PSYCHOLOGICAL SOCIETY, 1944

GENERAL MEETINGS

- March 30–April 4, 1944. ANNUAL GENERAL MEETING AT GLASGOW.
- March 30. **1st Session.** Business Meeting of the Society and Annual General Meetings of Sections.
- March 31. **2nd Session:**
 “The Psychology of ‘g’.” By C. SPEARMAN. (Read by Dr LL. Wynn Jones.)
3rd Session:
 “Some Observations on Performance Tests.” By HILDA W. OLDHAM.
4th Session:
 “The Applicability of the Terman-Merrill Tests in Scottish Children.” By D. KENNEDY FRASER.
 “Co-operation in Solving Problems.” By G. G. NEILL WRIGHT.
- April 1. **5th Session:**
 “Direct Projective Psycho-Therapy.” By MARGARET LOWENFELD.
 “Intellectual Deterioration associated with Hysterical Symptoms.” By E. J. G. BRADFORD.
6th Session:
 “Attitudes of Adults towards Property.” By J. A. WAITES.
 “Character Assessments from Handwriting.” By OLGA MARUM.
7th Session:
 Election of Officers and Council, followed by Presidential Address:
 “Social Differences in English Education.” By T. H. PEAR.
8th Session:
 “Perseveration and Musical Perception.” By MARTHA VIDOR.
- April 2. **9th Session:**
 “Colour Vision.” By R. W. PICKFORD.
10th Session:
 Open Session for Discussion.
11th Session:
 “Psychology and Public Relations.” By FREDERICK LAWS.
- April 3. **12th Session:**
 “Fluctuations in Mental Output.” By S. J. F. PHILPOTT.
 “The Method of ‘Shapes’ in Work Curve Analysis.” By F. W. WARBURTON.
 “Waves in Chemical Curves.” By H. E. GEORGIADIS.
13th Session:
 “Individual Differences in Fluctuation Rate.” By E. R. DENTON.
 “Factor Study of Decrement in Work Curves.” By D. F. WARD.
14th Session:
 “*Black Hamlet*—Some Meditations and Queries.” By JANE DARROCH.
 “Neurosis and Intelligence.” By H. J. EYSENCK.
 “*Gestalt* Theory and Thought.” By H. F. FRIEDLANDER.
15th Session:
 Open Session for Discussion.
- July 1, 1944. GENERAL MEETING IN LONDON.
 “A Study of Treason.” By A. M. MEERLOO.
- September 23, 1944. GENERAL MEETING IN LONDON.
 “Psychological Aspects of Moral and Social Progress.” By J. C. FLUGEL.
 Papers in comment by KARL MANNHEIM and R. H. THOULESS.
- October 16, 1944. EXTRAORDINARY GENERAL MEETING IN LONDON.
 Business:
 (1) To consider and vote on an application to form a Branch of the Society in Australia.
 (2) To elect new Australian Members.

SECTIONAL MEETINGS

MEDICAL SECTION

March 30, 1944. ANNUAL GENERAL MEETING AT GLASGOW.

EDUCATION SECTION

February 19, 1944. Discussion on the concept 'Intelligence Quotient':
 (1) The theoretical and statistical considerations. By P. E. VERNON.
 (2) Uses of the term by the professional psychologist. By J. M. GIBBS.

March 30, 1944. ANNUAL GENERAL MEETING AT GLASGOW.

INDUSTRIAL SECTION

March 30, 1944. ANNUAL GENERAL MEETING AT GLASGOW.
 June 15, 1944. "Some Aspects of Factory Inspection in War Time." By H. G. MAULE.
 August 29, 1944. "Proposed Work of the Medical Research Council's Unit of Applied Psychology at Cambridge." By K. J. W. CRAIK.
 October 10, 1944. "Personnel Selection in the Royal Canadian Navy." By E. S. W. BELYEA.
 November 28, 1944. "Some Problems of Selection and Training in War and in Peace." By E. A. BOTT.

SOCIAL PSYCHOLOGY SECTION

February 16, 1944. "Three Methods for Studying National Differences in Sense of Humour." By H. J. EYSENCK.
 "Some Difficulties in Participant Observation." By MARIE JAHODA.
 March 30, 1944. ANNUAL GENERAL MEETING AT GLASGOW.

BRANCH MEETINGS

February 5, 1944. (Edinburgh.) SCOTTISH BRANCH
 "Entroencephalography." By C. A. BEEVERS.
 "The Possible Contribution of Neurology to Psychology." By J. B. GAYLOR.
 March 4, 1944. (Glasgow.)
 "Conscience and Society." By RANYARD WEST.
 "Demonstration of a Schizophrenic Boy with a Phenomenal Memory for Dates." By F. BERLINER.
 "The Rorschach Test in Cases of Severe Concussional Head Injury." By O. L. ZANGWILL.
 November 4, 1944. (Edinburgh.)
 "The Noegenesis of Projection." By JOHN C. RAVEN.
 "Endopsychic Structure considered in terms of Object-Relationships." By W. R. D. FAIRBAIRN.
 December 2, 1944. (Glasgow.)
 "The Effects of Market Research on Press and Radio." By FREDERICK LAWS.
 "A Study of Infant Reading." By VIOLET M. McLAREN.
 January 8, 1944. (Birmingham.) MIDLAND BRANCH
 "The Sub-normal Child and the Residential School." By C. J. C. EARL.
 June 28, 1944. (Birmingham.)
 "The Testing, Educational Status and Training of a group of Adolescents in an Army Centre." By W. D. WALL.
 June 10, 1944. (Manchester.) NORTHERN BRANCH
 "Adolescent Leisure in a Lancashire Industrial Town." By W. HARPER.
 "A Proposal for Improving the Allocation of Filling-Factory Labour." By VALERIE C. STONE.
 "Growing Up Mentally." By T. H. PEAR.
 "The Development of Attitudes to Authority during Adolescence." By J. H. PRICE.
 July 29, 1944. (Leeds.)
 "What is a Drive?" By LL. WYNN JONES.
 "A Psychology of Freud's Unbelief." By W. E. SARGENT.
 "Some Non-Verbal Auditory Tests." By A. PEARSON.

BRITISH PSYCHOLOGICAL SOCIETY

REVENUE ACCOUNT AND BALANCE SHEET AS AT 31 DECEMBER 1944

Revenue Account for the year ended 31 December 1944

EXPENDITURE			INCOME		
1943	1943		1943		
£ s. d.	£ s. d.		£ s. d.	£ s. d.	£ s. d.
	<i>To</i> Rent	...	...	By Subscriptions 1941/43 ...	...
100 0 0	Rates	...	...	Subscriptions 1944 ...	...
118 6 3	Lighting, Heating and Cleaning	...	...		3 5 0
29 1 5	Salaries and State Insurance	...	...		1772 13 6
337 17 0	Printing and Duplicating ...	...	...		1775 18 6
95 11 0	Expenses of Meetings	...	...	Less Resignations, Removals and Sus-	
13 10 6	Traveling Expenses	...	...	pended Memberships ...	...
9 4 8	Postages and Telephone	...	...		87 11 6
84 4 8	Stationery	...	...	Transfer Fees:	
14 7 10	Insurance	...	...	Fellows (5) ...	7 17 6
17 1 3	Audit Fee 1944	...	...	Associates (9) ...	7 17 6
12 12 0	Teas	...	...	Interest on Investments ...	...
7 3 3	Sundry Expenses	...	...		15 15 0
9 13 10	<i>Journals, Publication Expenses and</i>	...	...		153 0 0
	<i>Reserves:</i>				
	280 15 9 General	...	...		
	304 6 8 Medical	...	...		
	130 13 0 Educational	...	...		
715 15 5	Extended General Meeting 1944 ...	...	...		
33 9 9	Scottish Branch Expenses	...	...		
7 7 8	Midland Branch Expenses	...	...		
1 19 7	Northern Branch Expenses	...	...		
	Committee of Professional Psychologists (Mental Health) ...	...	...		
	Bank Charges and Cheque Books ...	...	...		
1 17 1	<i>Depreciation:</i>				
	Library	...	...		
	Office Furniture and Fittings ...	...	...		
72 5 10					
1681 9 0					
	Balance, being Surplus of Income over Expenditure...	...	...		
£1681 9 0					£1857 2 0

Balance Sheet as at 31 December 1944

1943			1943			1943			1943		
£	s.	d.	£	s.	d.	£	s.	d.	£	s.	d.
22	8	2	136	17	4	35	6	0	219	14	3
30	8	6	1	15	6	58	7	0	3	10	3
122 0 0			4625 7 9			499 10 0			5125 7 9		
5571 2 11			162 0 0			5644 16 5			162 0 0		
73 13 6											